

# From Theory to Action

A pragmatic approach to  
digital preservation strategies and tools

June 30 and July 1, 2015

Co-Sponsored by: Northwest Archivists, Inc.  
Sustainable Heritage Network



Presenters:

Stacey Erdman

Jeff Hancks

Martin Kong

Danielle Spalenka



**Have one of the extras go around and do a quick test to make sure DA will run on the attendees' machines**

Hand out pretests as people walk in

# Logistics/Housekeeping

- Basic Logistics
- Handouts/Flash Drives
- Assessment/Evaluation (today and in 3 months) YOUR FEEDBACK IS VITAL
  - Pre-Test
  - Post-Test
  - Standard Workshop Evaluation
  - 3 Month Follow-up

## **First Up... The Day's Schedule!**

Pre and Post Test are only 5 questions. 3 month Follow up will likely be a short emailed survey, with a phone call if we don't hear anything back. We will be getting this contact info from you later in the day.

Pass out workshop evals now so attendees can jot down thoughts on what works and what needs improvement during the workshop

**\*\*GO OVER CONTENTS OF WORKSHOP PACKETS\*\***

## MORNING

### **Now – 9:45**

Collect Pre-tests  
Expected Outcomes  
Who we are & How we got here  
Levels of Preservation (*Activity*)

### **9:45 – 10:15**

Solution in Theory vs. Solution in Practice

*10:15 → Break*

### **10:30 – 11:00**

Your Pre-Ingest Workflow  
Accessioning a Collection (*Demo*)

### **11:00 – Noon**

Tools that POWRR investigated

## **LUNCH!**

## AFTERNOON

### **1:00 – 2:00**

Solution in Action: Accessioning a Collection (*Activity*)

### **2:00 – 2:30**

Assembling Your Team  
Your 3-3-3 Action Plan (*Activity*)

*2:30 → Break*

### **2:45 – 3:30**

Advocacy, Policy, Potential Solution Models

### **3:30 – 3:50**

Questions

### **3:50 – 4:00**

Post-Test

# Expected Outcomes

- You will understand that different digital preservation tools/services can perform different functions within the digital curation lifecycle, and be able to explain how these tools/services can be used within your institution's workflow.
- You will practice the initial pre-ingest steps necessary to accession a digital collection, as described in the OCLC report "Walk this Way," and gain the skills necessary to repeat this process at your institution.
- You will gain hands on experience with a basic digital preservation tool and understand how it can be used within your institution's workflow.
- You will take away resources that help align communication and advocacy, policymaking, and tool selection/implementation.
- You will create a 3-3-3 Action Plan to implement in the following 3 months that will move you closer to your digital preservation goals.

DANIELLE

## Who we are....and how we got here....

- Defining Moments → Found Some Friends
- Applied for Implementation Grant → Received a “Figure It Out” Grant → Received NEH grant

We’ve learned a lot...and are a lot like you!

Proud to be works-in-progress:



Northern Illinois University



WESTERN ILLINOIS UNIVERSITY



NATIONAL ENDOWMENT FOR THE  
humanities  
FIFTY YEARS



INSTITUTE of  
Museum and Library  
SERVICES

DANIELLE, then the rest

**At the end of this slide, have attendees introduce themselves: Name, position, and institution ONLY**

**\*\*See if Lynne has better slide\*\***

**ADD IMLS LOGO**

# Activity Time!

## 20 Minutes

### NDSA Levels of Preservation

Where can my institution place its Bingo chips?

- We'll go first
- Small Groups – Where do you think you fit in? (10 minutes)
- All Together – Poll of who is where!



**NATIONAL DIGITAL  
STEWARDSHIP ALLIANCE**

DANIELLE, then the rest of us

We still assign the Levels handout in advance, say our brief levels, turn them out into groups to start a discussion and get them familiar with each other, and then report out by show of hands?

NAME, INSTITUTION, ROLE ONLY!!!!!!!!!!!!!!



Table 1: Version 1 of the Levels of Digital Preservation

	Level 1 (Protect your data)	Level 2 (Know your data)	Level 3 (Monitor your data)	Level 4 (Repair your data)
Storage and Geographic Location	<ul style="list-style-type: none"> <li>- Two complete copies that are not collocated</li> <li>- For data on heterogeneous media (optical discs, hard drives, etc.) get the content off the medium and into your storage system</li> </ul>	<ul style="list-style-type: none"> <li>- At least three complete copies</li> <li>- At least one copy in a different geographic location</li> <li>- Document your storage system(s) and storage media and what you need to use them</li> </ul>	<ul style="list-style-type: none"> <li>- At least one copy in a geographic location with a different disaster threat</li> <li>- Obsolescence monitoring process for your storage system(s) and media</li> </ul>	<ul style="list-style-type: none"> <li>- At least three copies in geographic locations with different disaster threats</li> <li>- Have a comprehensive plan in place that will keep files and metadata on currently accessible media or systems</li> </ul>
File Fixity and Data Integrity	<ul style="list-style-type: none"> <li>- Check file fixity on ingest if it has been provided with the content</li> <li>- Create fixity info if it wasn't provided with the content</li> </ul>	<ul style="list-style-type: none"> <li>- Check fixity on all ingests</li> <li>- Use write-blockers when working with original media</li> <li>- Virus-check high risk content</li> </ul>	<ul style="list-style-type: none"> <li>- Check fixity of content at fixed intervals</li> <li>- Maintain logs of fixity info; supply audit on demand</li> <li>- Ability to detect corrupt data</li> <li>- Virus-check all content</li> </ul>	<ul style="list-style-type: none"> <li>- Check fixity of all content in response to specific events or activities</li> <li>- Ability to replace/repair corrupted data</li> <li>- Ensure no one person has write access to all copies</li> </ul>
Information Security	<ul style="list-style-type: none"> <li>- Identify who has read, write, move and delete authorization to individual files</li> <li>- Restrict who has those authorizations to individual files</li> </ul>	<ul style="list-style-type: none"> <li>- Document access restrictions for content</li> </ul>	<ul style="list-style-type: none"> <li>- Maintain logs of who performed what actions on files, including deletions and preservation actions</li> </ul>	<ul style="list-style-type: none"> <li>- Perform audit of logs</li> </ul>
Metadata	<ul style="list-style-type: none"> <li>- Inventory of content and its storage location</li> <li>- Ensure backup and non-collocation of inventory</li> </ul>	<ul style="list-style-type: none"> <li>- Store administrative metadata</li> <li>- Store transformative metadata and log events</li> </ul>	<ul style="list-style-type: none"> <li>- Store standard technical and descriptive metadata</li> </ul>	<ul style="list-style-type: none"> <li>- Store standard preservation metadata</li> </ul>
File Formats	<ul style="list-style-type: none"> <li>- When you can give input into the creation of digital files encourage use of a limited set of known open formats and codecs</li> </ul>	<ul style="list-style-type: none"> <li>- Inventory of file formats in use</li> </ul>	<ul style="list-style-type: none"> <li>- Monitor file format obsolescence issues</li> </ul>	<ul style="list-style-type: none"> <li>- Perform format migrations, emulation and similar activities as needed</li> </ul>

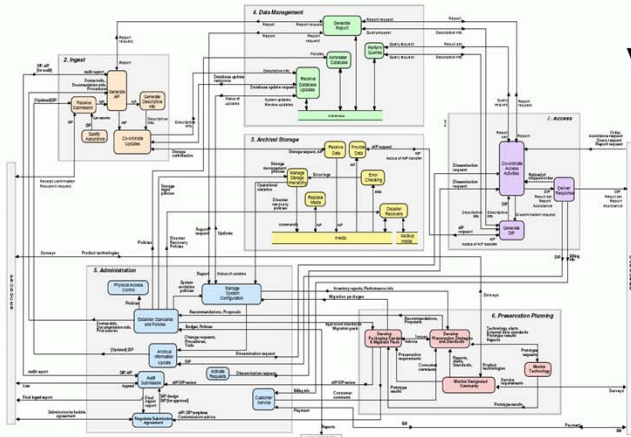
<http://www.digitalpreservation.gov/ndsa/activities/levels.html>

DANIELLE Mention that this is version 1...work in progress!!!

**Meg's note: make clear interested in the "bulk" of capabilities—will be shorter discussion if we ask people to just state Level 1, 2, 3, or 4 rather than everyone sharing details on which specific cells they can do.**

# So.....How do we get from here to there?

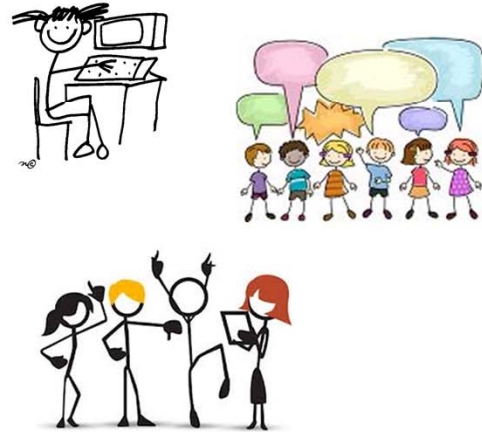
## Solution in Theory



Scary OAIS Spaghetti Monster

vs.

## Solution in Practice



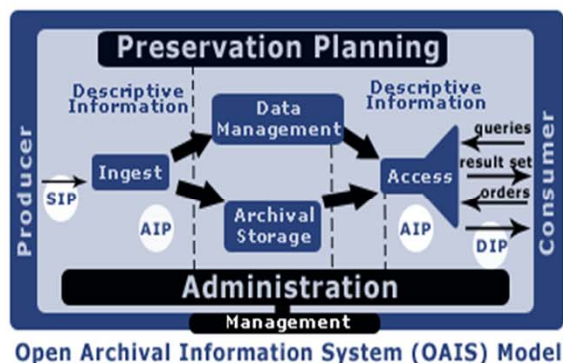
Solution in Theory: Scary OAIS Spaghetti Monster

Solution in Practice: Figure out what you have, start talking to people, build a team and eventually a policy!



# Solution in Theory

- OAIS (Open Archival Information Systems) and other schematic models
- TRAC Certification (Trustworthy Repositories Audit & Certification)
- TDR ISO 16363 (Trustworthy Digital Repository ISO Standard)
- Curation Lifecycles that don't look a thing like our current workflows



*SIPs, AIPs, DIPs, Oh my!*

**ACTUALLY GO OVER WHAT AIPS SIPs AND DIPs ARE..>USE THE SIMPLIFIED LANGUAGE WE PUT IN THE WHITE PAPER**

**SIP** – Ingest (Accessioning)

**AIP** – Data Management (Storage)

**DIP** - Access

In just over 10 years, **there are 5 repositories that have been able to become certified!!!** ([Chronopolis Report](#); [Hathitrust Report](#); [Portico Report](#); [Scholars Portal](#); [CLOCKKS](#), which received the highest score of any org) OAIS was conceived in 1996 and accepted as an ISO standard in 2002.

**A Note:** These are all valuable things that benefit the field of digital preservation greatly.... We just don't want you to become overwhelmed by them and grind to a halt before you take your first steps...like we did!

In the intro (to the 90 page document called the TRAC Criteria & Checklist," the process came out of an identified need...

"The [OAIS] reference model (ISO 14721) provides a common conceptual framework describing the environment, functional components, and information objects within a system responsible for the long-term preservation of digital materials. Long before it became an approved standard in 2002, many in the cultural heritage community had adopted OAIS as a model to better understand what would be needed from digital preservation systems.

"Institutions began to declare themselves 'OAIS-compliant' to underscore the trustworthiness of their digital repositories, but there was no established understanding of

'OAIS-compliance' beyond meeting the high-level responsibilities defined by the standard. There were certainly no criteria for measuring compliance."

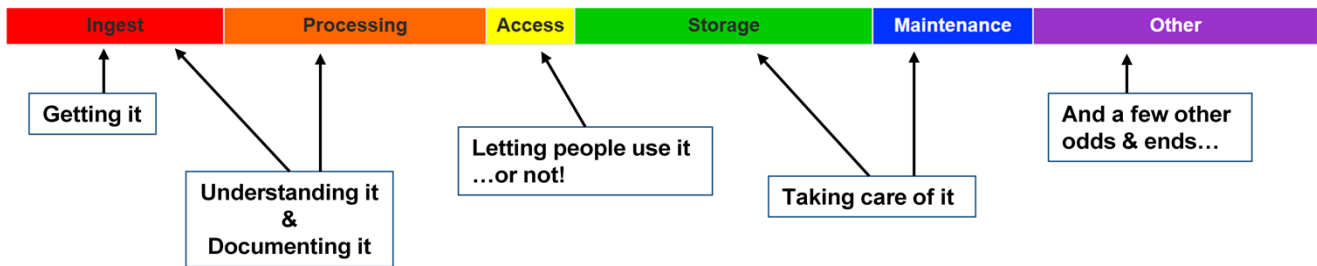
## **We can be *ready*.**

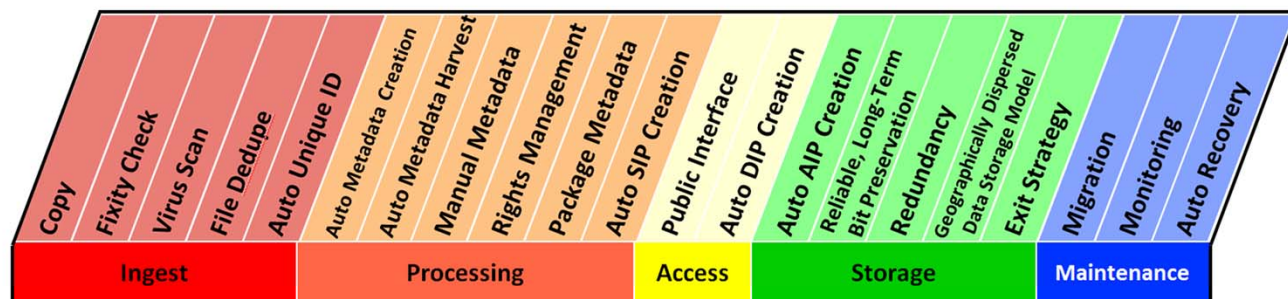
- We can intellectually map how digital preservation works to how we work.
- We can upgrade our metadata and recordkeeping practices for the next steps.
- We can triage our data for ingest.
- We can build policies and plans, which in turn help us choose tools.
- We can better educate ourselves, our stakeholders, and our funders.

# Solution in Practice

AKA Good Enough DP for real people!!

Our take on what you need to consider when thinking about your digital stuff.....





Our take on some things that need to happen or be considered along the way to this “*Digital Preservation*” thing....

**\*\*Show the handout! There is a handout in the packets that show our definitions for all of these.\*\***

# Solution in Practice

AKA Good Enough DP for real people!!

Some things to keep in mind.....

- Not all tools and services are created equal.
  - Some tools/services do specific tasks (*microservices*).
  - Some tools/services combine multiple microservices (*you guessed it....macroservices!*).
- Starting small is good enough!
- Knowing what you have is crucial.
  - Write. It. Down. And maintain it.

\*\* Monologue to talk them off the edge.\*\* You don't have to pick the tools/services that will do ALL of the steps from the get go.

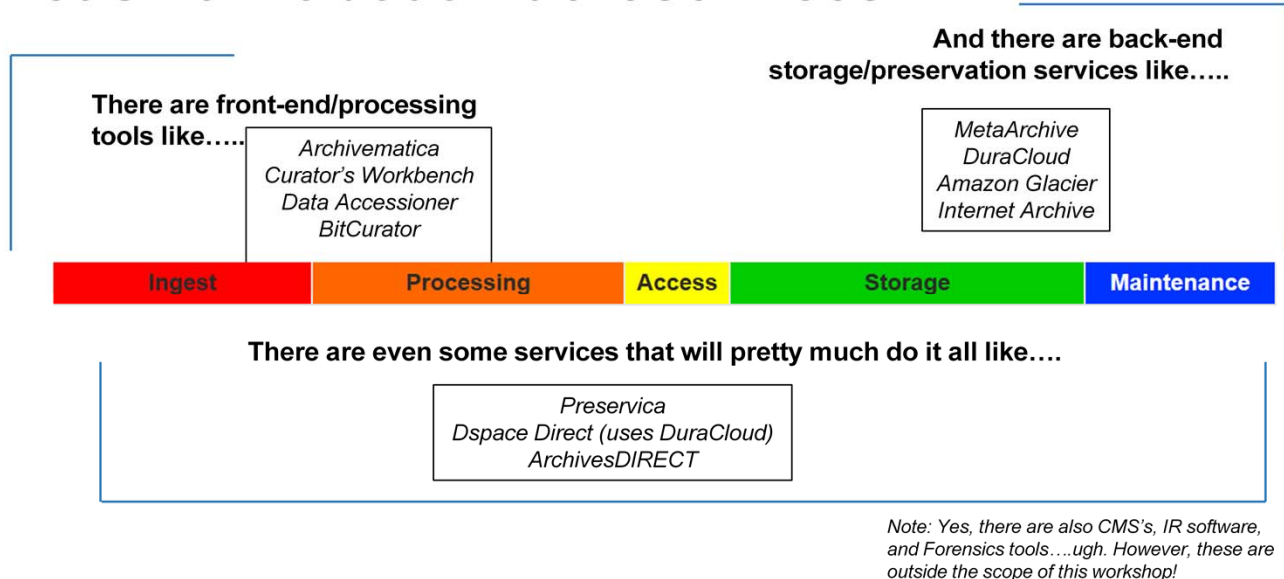
Start with a step or two. And it doesn't even have to be with a fancy schmancy tool!

Starting with a simple microservice tool will get you closer to your goals AND you can use them NOW!

- Baby steps still move you forward..... See "Walk This Way".

CAVEAT: Time Sensitive. Software changes quickly. Like, REALLY QUICKLY. Reference COPTR

# Let's Talk about Macroservices....



AND you have to figure out what works best with what!!! But we have done some of that so you don't have to!

Having a basic knowledge of DP...they should already know the difference between storage and preservation. But it may come up as a question anyway.

Explain what we mean by processing and backend. Define difference between DP and IR. Details on these specific tools later today. WE PROMISE.

**NOTE ABOUT BITCURATOR:** The BitCurator project proposed to build a software environment that would integrate digital forensics tools and methods into the workflows and collection management environments of libraries and archives, and that would support properly mediated public access to forensically acquired data. Using digital forensics approaches can help us ensure the authenticity, integrity, and provenance of digital materials.

The project was a joint effort between the School of Information and Library Science at the University of North Carolina-Chapel Hill and the Maryland Institute for Technology in the Humanities, and was funded by an Andrew W. Mellon grant between 2011-2014. They produced the BitCurator software environment, which is built on a stack of free and open source digital forensics tools and associated software libraries. BitCurator can be installed in a Linux environment, or run as a Virtual machine on top of most contemporary operating systems. **Features of BitCurator include: pre-imaging data triage, forensic disk imaging, file system analysis and reporting, identification of private and individually identifying information, and export of technical and other metadata**

# Clarification: Preservation vs. Access

## Long term access (Preservation)

- **Purpose:** ensure long-term access
- **Focus:** current & future users
- Relies on **proven (reliable)** technologies to preserve digital objects across generations of technology
- **Accumulates** metadata over the life cycle to trace preserved content
- Preservation systems **create** new versions of digital objects for access to deliver as needs change over time

## Short term access

- **Purpose:** provide content to users now
- **Focus:** current
- Relies on **cutting edge** technologies to provide best and fastest access at a point in time
- **Selects** metadata needed to use and understand content
- Access systems **deliver** objects with user-oriented services



## REALLY HONE IN ON ACCESS VS. PRESERVATION

From Library of Congress: We are focused on preparing content for preservation. Access is currently gravy.



**BREAK TIME!**

Back by 10:30, please

Next up: Your Pre-Ingest Workflow

## Actual Conversation, ca. 2004

“I’d like our institution to be the home for your literary papers.”

\*gets handed flash drive\*



HEY! We digitized this collection and it's on the server! Now what?  
Has anyone heard this question: Hey, do you want this flash drive for your collection/archive?  
The answer: "suuuuuurrrre."

# Don't Panic - Your Pre-Ingest Workflow

aka Wrangling your digital stuff before you can get it into a shiny system

**NOTE: This is only ONE way to do this... Everyone's workflow is a little different!**

*Hang tight for a demo  
and some hands-on  
practice of this!*

## Starting from scratch:

- Begin an Inventory Spreadsheet
- Run accessioning tools (creates basic preservation metadata files in XML for you!)
  - Move everything to a stable carrier (like a network drive)
- Make an Access Copy from your Master Copy
- Continue populating Inventory Spreadsheet (if needed)
- OPTIONAL: Keep original media
  - ✓ *Most of these will cost you more time than money*
  - ✓ *Document what you do pre-ingest. For future you.*
  - ✓ *Remember: Good enough is just fine. For now.*

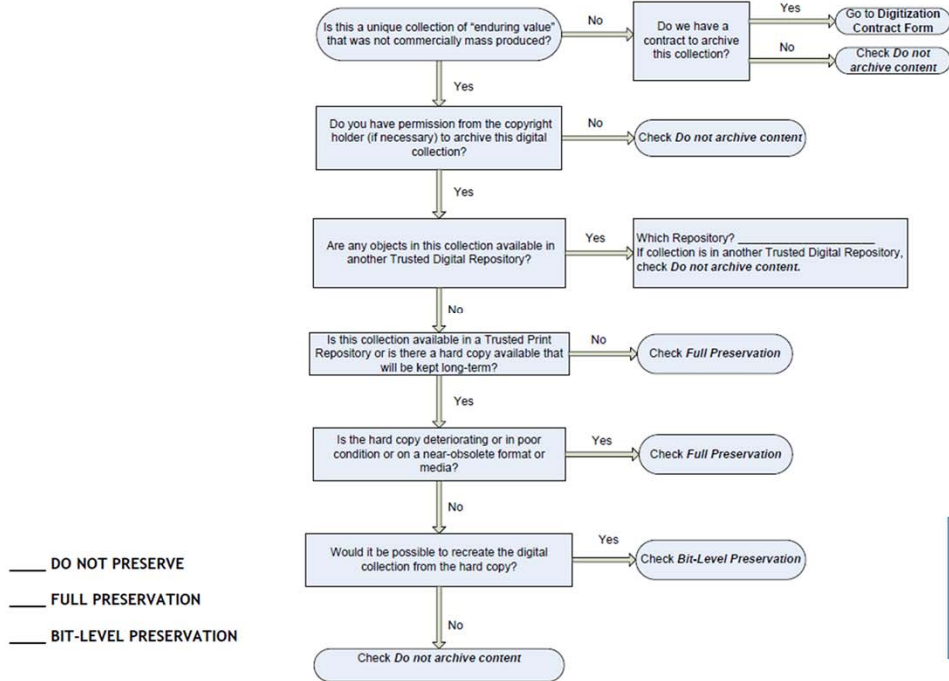
Good enough = one way to do this

It is important to begin somewhere. As an archivist, these are the same steps as what you would do for analog material:

- 1.) Know what you have (inventory). Can be as simple as creating a spreadsheet
- 2.) Run accessioning tools and move to stable carrier (these tools help tell you what you have)
- 3.) **MAKE ACCESS COPY**
- 4.) Continue inventory

Consider using a clean machine and/or writeblocker

**DIGITAL PRESERVATION DECISION FLOWCHART**



**Courtesy of:**  
 Tawnya Keller, *Digital Preservation Archivist*  
 University of Utah

\*\*Note that this is in their handout and has a side 2 to help them articulate different decision points.\*\*

# Pre-Ingest Inventory Spreadsheet Categories

These suggestions follow the recommended DPOE step “Identify” as locally defined by curator/archivist. Example at: [http://www.carli.illinois.edu/sites/files/digital\\_collections/documentation/digipres\\_identify.pdf](http://www.carli.illinois.edu/sites/files/digital_collections/documentation/digipres_identify.pdf)

- Category (digitization project; born digital; university archives)
- Title and Description
- Date(s) (date range of what’s IN there or date of creation if born digital)
- Location (CD, Jump drive, server location?)
- Extent (quantity: 48 journal issues; 106 images; 2 TB of video)
- Format (file formats: PDF, .Jpeg, Animated GIF, Wordstar2.0 file)

This is YOUR inventory... YOU get to decide if it needs additional fields, if some can be deleted, etc. You are the boss of this!

Category	Title and Description	Date	Location	Extent	Format
----------	-----------------------	------	----------	--------	--------

**FILL OUT WHAT YOU CAN AS YOU WOULD WITH ANY NORMAL ACCESSION**

DPOE is a Library of Congress Digital Preservation and Outreach Program  
<http://www.digitalpreservation.gov/education/>



This is just ONE way of starting it – simply filling out an spreadsheet to keep track of what you have.

These are basic categories to help identify – knowing what you have is a great first step that you can do now!

**\*\*Tell story of example donor\*\***

What we know of the collection: “Jane and John Moneybags, Class of 2006, have kept materials related to research and patent for a feline health product. The collection will be called “A Curator’s Cat Collection.”

Category	Title & Description	Date	Location	Extent	Format
(locally defined; project name? content creation method?)	(Donor applied and/or yours... what's your local practice?)	(YYYYMMDD or other locally defined format for accession date)	(Storage place of choice -- networked server recommended)	(Quantity of folders, files, by type or total size)	(What extensions are involved: .jpg, .tif, .xls?)
Special Collections, mixed; digitized and born digital	A Curator's Cat Collection. Donated by Jane (nee Pennypincher) and John Moneybags, Class of 2006. Feline Health Research. No restrictions on access; some material may have copyright restrictions by law	20150424	C:\Users\User\Desktop\NewAccession\Masters		

This is what the spreadsheet looks like with our example collection when we fill it in on our own.

Notice that we have left the "Extent" and "Format" blank. We'll show you a tool that will help fill in this information.

# Data Accessioner

Digital POWRR Tool Evaluation Grid	Ingest			Processing						Access		Storage				Maintenance			Other					
	Copy	Fixity Check	Virus Scan	File Dedupe	Auto Unique ID	Auto Metadata Creation	Auto Metadata Harvest	Manual Metadata	Rights Management	Package Metadata	Auto SIP Creation	Public Interface	Auto DIP Creation	Auto AIP Creation	Reliable, Long-Term Bit Preservation	Redundancy	Geographically Dispersed Data Storage Model	Exit Strategy	Migration	Monitoring	Auto Recovery	Open Source	Clear Documentation	Cost
Duke Data Accessioner	x	x			x	x	x	x		x	x											x		Free

## Explain why we are demoing Data Accessioner

Go through this slowly.... DA was built out of the need for a simple interface to allow technical services staff to have an easy way to quickly migrate data off disks and other external media onto a file server for basic preservation, further appraisal, arrangement and description. It also provides an easy way to integrate common metadata creating tools at the time of migration rather than after the fact. These tools include a MD5 checksum generator, JHOVE (Jstor/Harvard Object validation Environment) and DROID (Digital Record Object Identification). JHOVE provides functions to perform format-specific identification, validation, and characterization of digital objects. DROID is designed to meet the fundamental requirement of any digital repository to be able to identify the precise format of all stored digital objects, and to link that identification to a central registry of technical information about that format and its dependencies.

With a simplified interface, and written in Java (which is platform independent), it is intended to be easily adopted by smaller institutions with little or not IT Staff support

The tool is composed of several components bound into a single graphical interface. It recursively navigates a file tree and creates a copy of the tree in a given destination with the option of skipping specified files and directories. It creates a MD5 checksum for the file before copying it, and then, after copying creates another on the new copy and compares the two. If there is any difference, it creates an error and will notify the user. It also sets the last modified date of the copy to that of the original.

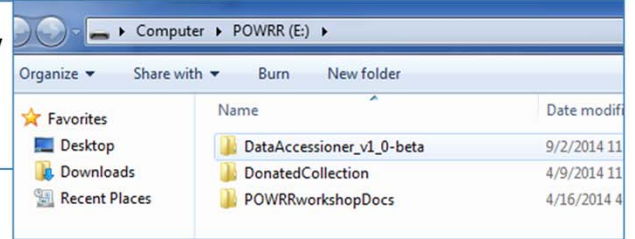
Once the migrator has started, it instructs the adapter plugins to do their work and to send

the results to the metadata manager for output in XML.

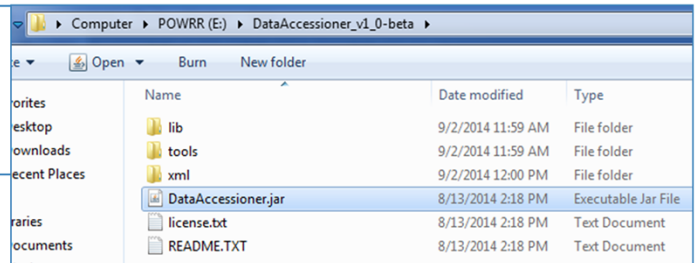


## 1. Insert flash drive and open the explorer window

Data Accessioner  
Donated Collection Folder  
Other stuff.....



## 2. Navigate to DataAccessioner.jar and open it



Walk through opening DataAccessioner → DataAccessioner.jar

Switch to live Data Accessioner demo

# XML Output

```
<?xml version="1.0" encoding="UTF-8"?>
<collection name="Curator's Cat Collection" xmlns="http://dataaccessioner.org/schema/dda-0-3-1">
  - <accession number="2014-Sept-02">
    - <folder name="A Curator's Cat Collection" last_modified="2014-04-09T11:59:08.000">
      - <dcx:description xmlns:dcx="http://purl.org/dc/xml/">
        <dc:description xmlns:dc="http://purl.org/dc/elements/1.1/">A quirky collection of kitties</dc:description>
        <dc:source xmlns:dc="http://purl.org/dc/elements/1.1/">Wealthy Donor x</dc:source>
        <dc:title xmlns:dc="http://purl.org/dc/elements/1.1/">A Curator's Cat Collection</dc:title>
      </dcx:description>
    - <folder name="Stacey's Kitties" last_modified="2014-04-09T11:59:08.000">
      - <file name="10154101_10103012097484377_1732212283_n.jpg" last_modified="2014-04-02T17:21:50.000">
        - <premis:object xsi:type="premis:file" xmlns:uuid="java:java.util.UUID" xmlns:fits="http://hul.harvard.edu
          xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xmlns:premis="info:ic/xmlns/premis-v2">
          - <premis:objectIdentifier>
            <premis:objectIdentifierType>uuid</premis:objectIdentifierType>
            <premis:objectIdentifierValue>5096d91f-407c-40db-8362-ae2d784b26f2</premis:objectIdentifierVal
            </premis:objectIdentifier>
          - <premis:objectCharacteristics>
            <premis:compositionLevel>0</premis:compositionLevel>
            - <premis:fixity>
              <premis:messageDigestAlgorithm>MD5</premis:messageDigestAlgorithm>
              <premis:messageDigest>3ab9f14fec4526582c1686e36e4b4706</premis:messageDigest>
              <premis:messageDigestOriginator>OIS File Information</premis:messageDigestOriginator>
            </premis:fixity>
            <premis:size>80112</premis:size>
          - <premis:format>
            - <premis:formatDesignation>
              <premis:formatName>JPEG File Interchange Format</premis:formatName>
              <premis:formatVersion>1.01</premis:formatVersion>
            </premis:formatDesignation>
            <premis:formatNote>Identified by: Exiftool v9.13</premis:formatNote>
          </premis:format>
          - <premis:format>
            - <premis:formatDesignation>
              <premis:formatName>JPEG image data</premis:formatName>
            </premis:formatDesignation>
            <premis:formatNote>Identified by: file utility v5.03</premis:formatNote>
          </premis:format>
        </premis:object>
      </file>
    </folder>
  </accession>
</collection>
```

Basic descriptive and Dublin Core metadata that you created

Extracted metadata:  
tree hierarchy of your  
accession  
(folder names, file  
names, last modified,  
size, and more!)

Fixity/Checksum: MD5

Identifying  
Information  
(Exiftool)

Identifying  
Information  
(file utility)

Here is what the XML File tells us:

- Basic descriptive metadata
- Extracted metadata (folder tree hierarchy)
- Fixity/checksum: MD5
- Identifying Information (Exiftool)
- Identifying Information (File Utility)

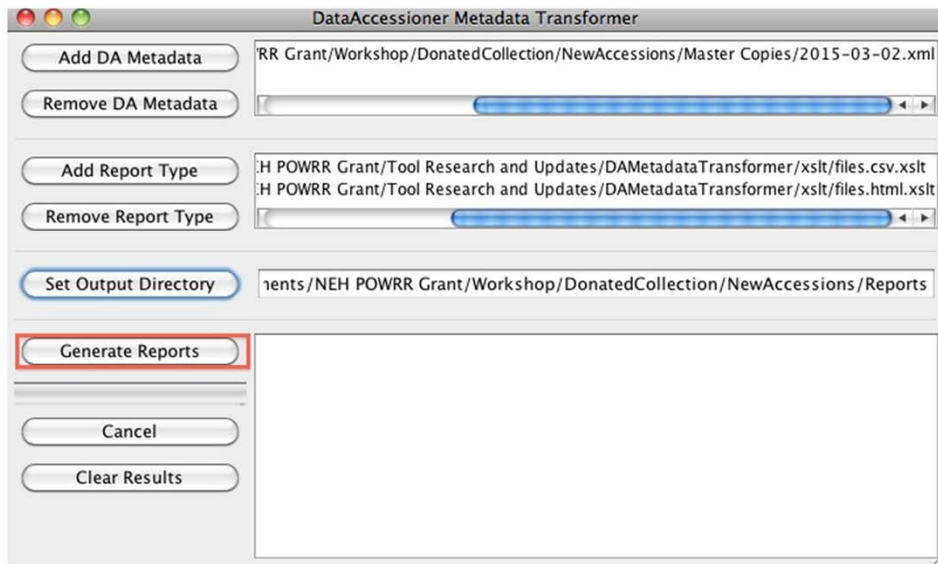
At end of product overview, mention that XML is a standard that many (all?) systems we tested can understand with little human intervention.

NotePad ++ is a free tool that can help you browse through and search the xml file, perform counts, etc.

But hold on! Isn't there a better way to view this raw XML file?

There IS! The developer of DA has developed a way to transform this raw XML into a nice report

# DataAccessioner: Metadata Transformer

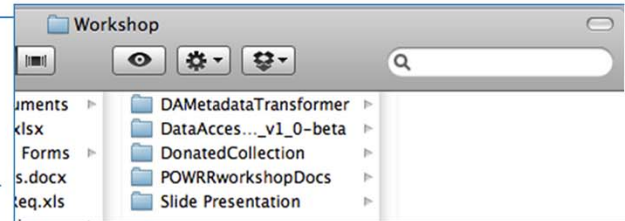


We would like to give a huge shout out to Seth Shaw, who not only updated Data Accessioner but added this wonderful reporting tool. Seth Shaw is amazing and nice. This reporting tool takes the raw XML and creates lovely reports that you can also attach to your collections (or eventually upload to a shiny system!)

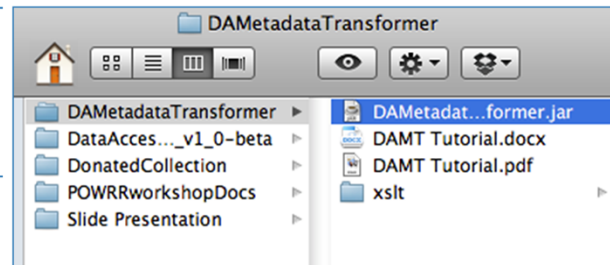
- CSV File – Gives you a high level view of the collection
- HTML – allows you to see the the details in a more palliative view, seeing the metadata you added in DA.

**1. Open the Explorer Window and open DAMetadadataTransformer Folder**

Data Accessioner  
Donated Collection Folder  
DAMetadadataTransformer  
Others...



**2. Navigate to DAMetadadataTransformer.jar and open it**



Walk through opening DAMetadadataTransformer → DAMetadadataTransformer.jar

## Switch to live DA:MT Reporting Tool demo

There will be a PDF included in the packets that includes step-by-step instructions. This tool is just to help see the data in HTML and CSV formats so you don't have to look at the raw XML.

Category	Title & Description	Date	Location	Extent	Format
(locally defined; project name? content creation method?)	(Donor applied and/or yours... what's your local practice?)	(YYYYMMDD or other locally defined format for accession date)	(Storage place of choice -- networked server recommended)	(Quantity of folders, files, by type or total size)	(What extensions are involved: .jpg, .tif, .xls?)
Special Collections, mixed; digitized and born digital	A Curator's Cat Collection. Donated by Jane (nee Pennypincher) and John Moneybags, Class of 2006. Feline Health Research. No restrictions on access; some material may have copyright restrictions by law	20150424	C:\Users\User\Desktop\NewAccession\Masters	38.83 MB	11 jpg, 1 pdf, 3 PDF/A, 2 MPEG v4

Look at all the data you are able to fill in! This is just ONE way of doing it. This allows you to get started immediately.

# Congratulations!

You just did the first few steps in the digital curation lifecycle.

Digital POWRR Tool Evaluation Grid	Ingest			Processing						Access		Storage				Maintenance			Other					
	Copy	Fixity Check	Virus Scan	File Dedupe	Auto Unique ID	Auto Metadata Creation	Auto Metadata Harvest	Manual Metadata	Rights Management	Package Metadata	Auto SIP Creation	Public Interface	Auto DIP Creation	Auto AIP Creation	Reliable, Long-Term Bit Preservation	Redundancy	Geographically Dispersed Data Storage Models	Exit Strategy	Migration	Monitoring	Auto Recovery	Open Source	Clear Documentation	Cost
Duke Data Accessioner	x	x			x	x	x			x	x											x		Free

Well, we did it. *Your turn* comes after lunch!

But first..... Macroservices! WOO!



# Macroservices: Doing it all! Sort of.

Using simple tools, like Data Accessioner, is what you can do while you are petitioning your institution for a more robust solution like...

- Archivematica
- Curator's Workbench
- DuraCloud
- MetaArchive
- Preservica
- Internet Archive

**Please Keep In Mind...**

This is NOT exhaustive

Software changes quickly!

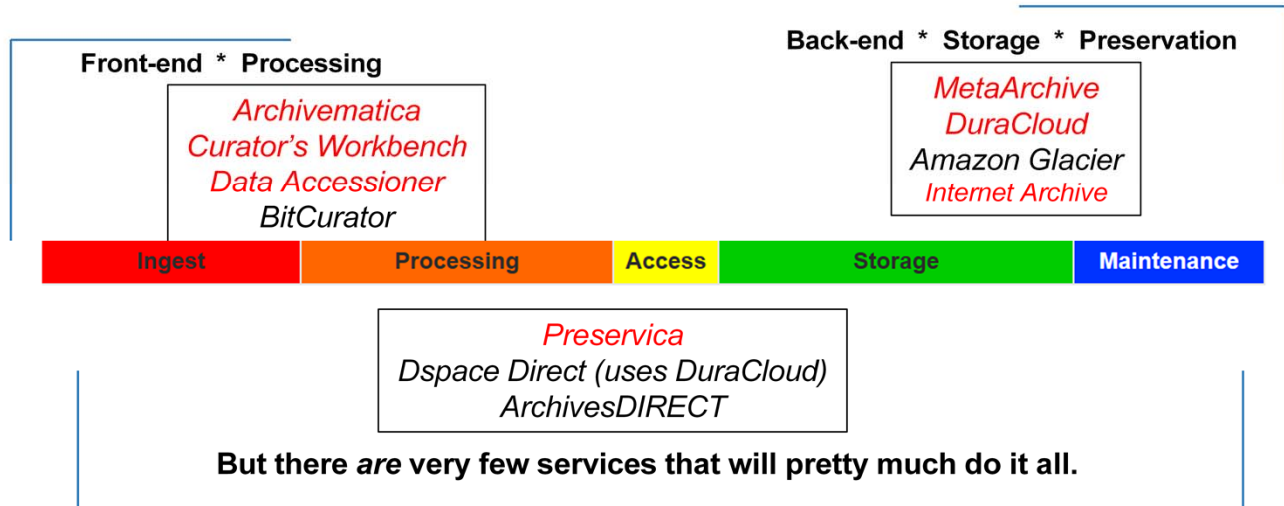
Based on availability at time of testing and our perceived needs

These are tools we tested in 2013 as part of the IMLS-phase. Again, not an exhaustive list. **WE DO NOT RECOMMEND OR ENDORSE ANY ONE TOOL/SERVICE.** This list was meant to be a sample of tools/services that perform certain actions. The list was recommended by our board of advisors during the IMLS-phase.

# Remember this?

Most tools and services only perform *some* of the functions in a digital curation lifecycle.

*\*Tools/Services in RED were tested in-depth by POWRR*



Jaime note: To me, Front-end means the initial steps in the Dig Curation lifecycle, and not the "Access" part of things. Perhaps we can make it clearer by clarifying what it is for the archivist/curator facing and what would be end-user facing (access/DIP/etc.)

**We can also talk about which phases spit out SIP's DIP's and AIP's and where those are used subsequently and by whom.**

Most of these are Open Source or FREE tools.

## A note about the word “free”



NOT



Open source software requires resources to install, maintain, and improve it.

Some tools are free open source tools. But it's important to note that open source tools are free like kitties, and not like beer. There's A LOT of work and maintenance that goes along with that free kitty, whereas a free beer allows you to sit back and enjoy. Open source software requires resources to install, maintain, and improve it.

# Front-end/Processing: Curator's Workbench

Digital POWRR Tool Evaluation Grid	Ingest			Processing					Access		Storage				Maintenance			Other						
	Copy	Fixity Check	Virus Scan	File Dedupe	Auto Unique ID	Auto Metadata Creation	Auto Metadata Harvest	Manual Metadata	Rights Management	Package Metadata	Auto SIP Creation	Public Interface	Auto DIP Creation	Auto AIP Creation	Reliable, Long-Term Bit Preservation	Redundancy	Geographically Dispersed Data Storage Model	Exit Strategy	Migration	Monitoring	Auto Recovery	Open Source	Clear Documentation	Cost
**Curator's Workbench	x	x			x	x	x			x	x											x		Free

**\*\*Take time to explain what we were looking at when we tested the tool. Explain the process of testing/rubric we used to evaluate\*\***

**Open source**

Briefly discuss this is a front end processing tool similar to DA, only is a MODS-based platform.

# Front-end/Processing: Archivematica

Digital POWRR Tool Evaluation Grid	Ingest			Processing					Access		Storage				Maintenance			Other					
	Copy	Fixity Check	Virus Scan	File Dedupe	Auto Unique ID	Auto Metadata Creation	Auto Metadata Harvest	Manual Metadata	Rights Management	Package Metadata	Auto SIP Creation	Public Interface	Auto DIP Creation	Auto AIP Creation	Reliable, Long-Term Bit Preservation	Redundancy	Geographically Dispersed Data Storage Model	Exit Strategy	Migration	Monitoring	Auto Recovery	Open Source	Clear Documentation
**Archivematica	X	X	X		X	X	X	X	X	X		X	X					X			X	X	Free

We tested version 0.9 and 1.0

Notice this does a lot of the front-end processing: Ingest, Processing, with some maintenance.

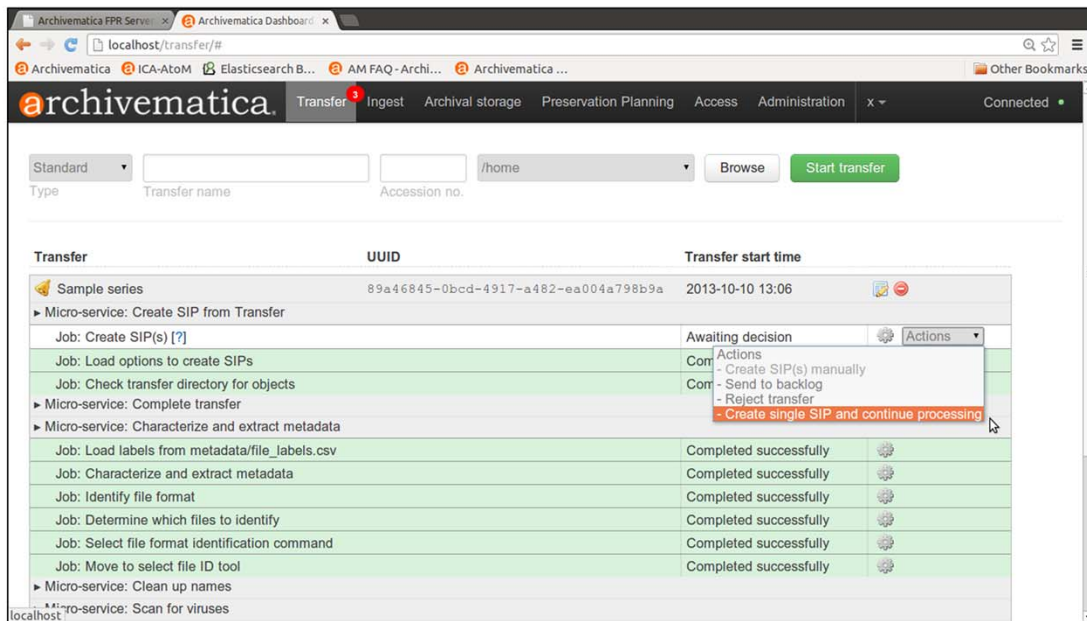
# Front-end/Processing: Archivemata

- Open source/free software
- Requires IT support and administration (Virtual Machine, Ubuntu Server, etc.)
- Microservices run by themselves
- Shows all the steps for AIP, SIP, DIP
- Capability to upload own metadata
- Errors stop everything
- Great Google users group support
- Integrates with Content DM & DSpace
- Bundled with ICA-AToM (archival content management system like ARCHON)
- Hosted version now available
- File transfers not intuitive
- Slower processing, but that could be due to the fact that we are used to desktop-based applications

Hosted version now available – ArchivesDirect (will talk momentarily)  
Archivemata 1.4 and Storage Service 0.7 just released May 27, 2015. This version includes a number of bug fixes, and new tools/features. Highlights include:

- CONTENTdm integration for hosted or non-hosted environments
- Dspace integration, including capture of parent-child relationships
- Siegfried file identification tool
- Islandora/Archidora plugin
- Recovery of a backed up AIP

# Archivematica: Transfer Collection



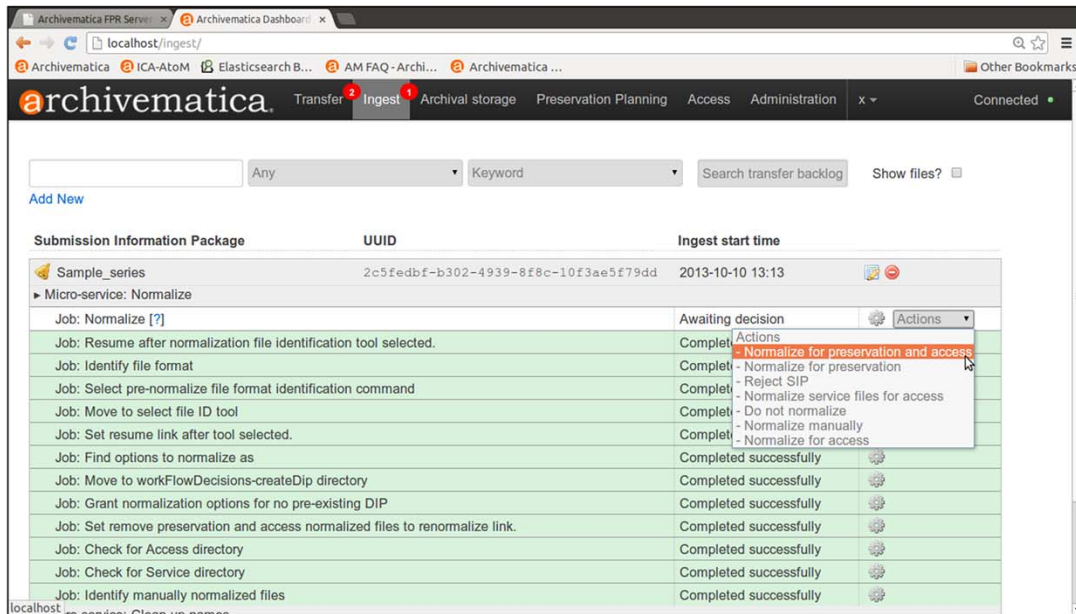
In Archivematica, Transfer is the process of transforming any set of digital objects and/or directories into a SIP.

In the Transfer tab of the Dashboard, the user moves digital objects from source directories accessible via the Storage Service into Archivematica.

Once uploaded to the dashboard, transfers run through several micro-services: UUID assignment; checksum verification (if checksums are present); package extraction (i.e. unzipping of zipped or otherwise packaged files); virus checking; indexing; format identification and validation; and metadata extraction.

At the end of transfer, the user creates a SIP from one or more standard transfer(s). Once this is done, the SIP is moved into ingest.

# Archivematica: Normalization On Ingest



During ingest, digital objects are packaged into SIPs and run through several micro-services, including normalization, packaging into an AIP and generation of a DIP. (this screenshot is of Normalization) (ie WordPerfect to Word)



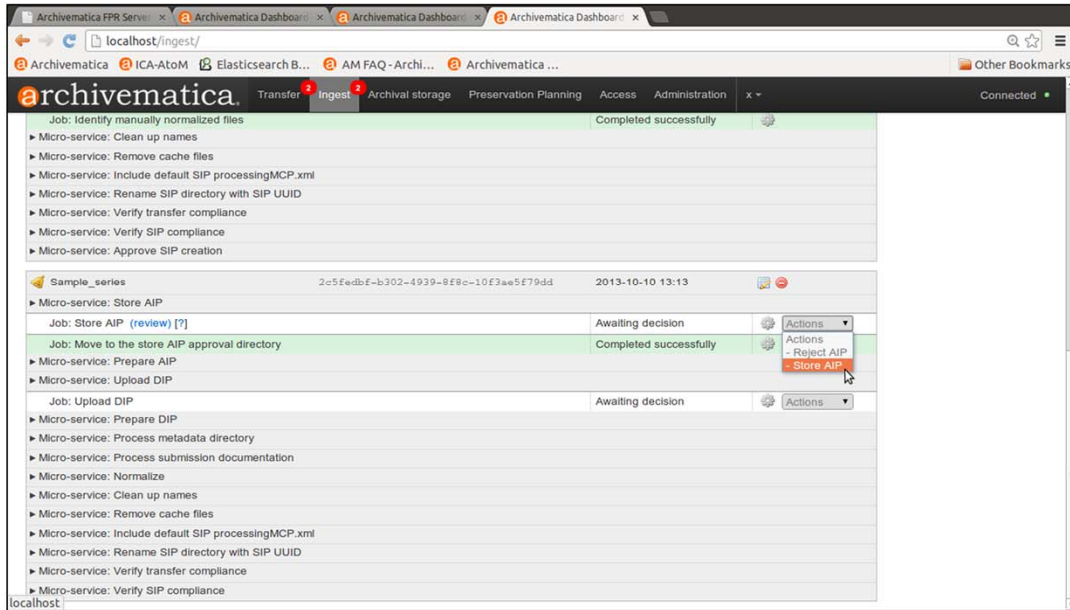
# Archivematica: Add Metadata

The screenshot shows the Archivematica web interface for adding metadata. The breadcrumb trail is 'Ingest / Test\_files / Metadata / Add'. The main heading is 'Metadata' with a sub-heading 'Test\_files'. The form includes the following fields:

- Applies to:** A dropdown menu with 'Test\_files' selected. A note below reads: 'Metadata can be added at the SIP/AIP level only'.
- Title:** Text input field containing 'Test files'.
- Creator:** Text input field containing 'Clancy King'.
- Subject:** Text input field.
- Description:** Text area.
- Publisher:** Text input field.
- Contributor:** Text input field.
- Date:** Text input field containing '2012/04/05'. A tooltip below reads: 'Use ISO 8061 (YYYY-MM-DD or YYYY-MM-DD/YYYY-MM-DD)'. A larger tooltip on the right side of the form reads: 'A point or period of time associated with an event in the lifecycle of resource. (ISO 15838)'.

This easy to use form allows you to create metadata, either local metadata or using a standard (METS, MODS, DC, etc)

# Archivematica: Add AIP to Storage



At this point, you can store the AIP to a storage system (your own server, another storage solution, etc)

# Back-end/Preservation: DuraCloud

Digital POWRR Tool Evaluation Grid	Ingest				Processing						Access		Storage				Maintenance			Other			
	Copy	Fixity Check	Virus Scan	File Dedupe	Auto Unique ID	Auto Metadata Creation	Auto Metadata Harvest	Manual Metadata	Rights Management	Package Metadata	Auto SIP Creation	Public Interface	Auto DIP Creation	Auto AIP Creation	Reliable, Long Term Bit Preservation	Redundancy	Geographically Dispersed Data Storage Model	Exit Strategy	Migration	Monitoring	Auto Recovery	Open Source	Clean Documentation
**DuraCloud	x	x		x	x	x	x	x				x		x	x	x	x		x	x	x	x	Varies

# Back-end/Preservation: DuraCloud

- Nonprofit; Open Pricing; Community buy-in
- Cloud storage/preservation solution
- Different storage provider options
- Hosted service (requires little to no IT support on your end!)
- Some microservices available (like health checks that verify checksums)
- Different options/methods for uploading content (bulk, single item, etc.)
- Intuitive uploads and file management
- Easy exit strategy
- Easy integration with DSpace
- New: Integrated with hosted version of Dspace
- Media streaming based on Amazon's Cloud service
- Responsive customer service with very good documentation
- Affordable; Scalable; Easy to get started

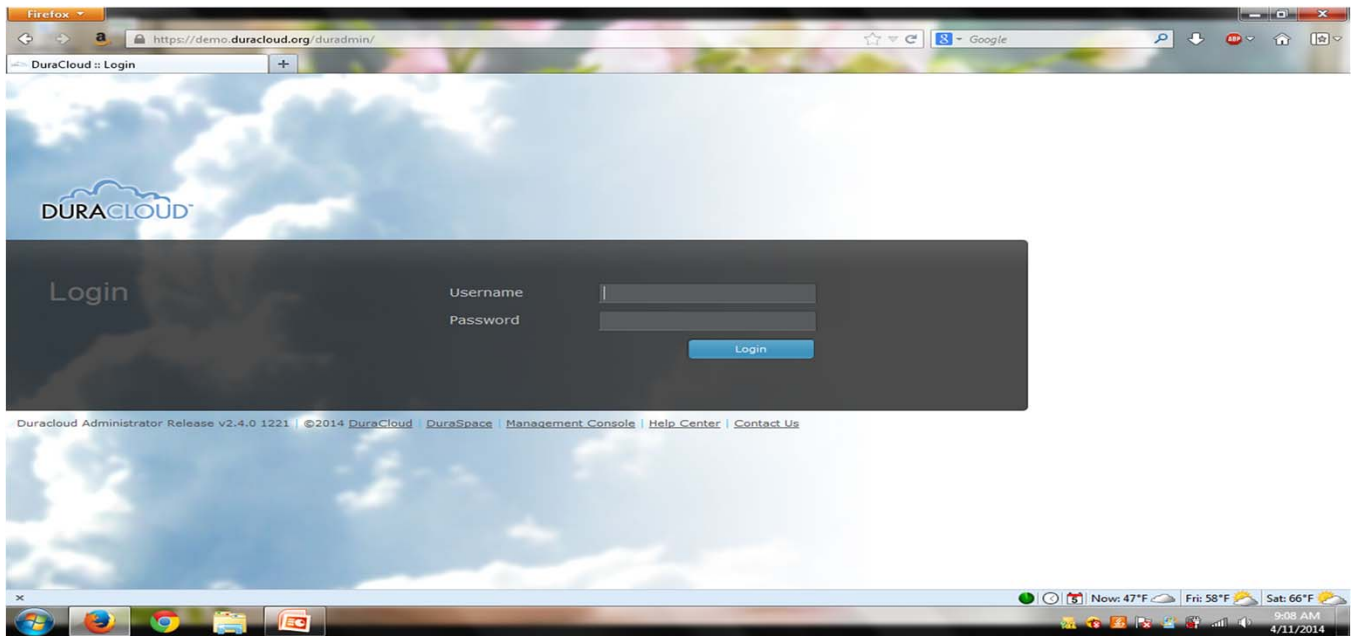
# DuraCloud.org

Head to the website for...

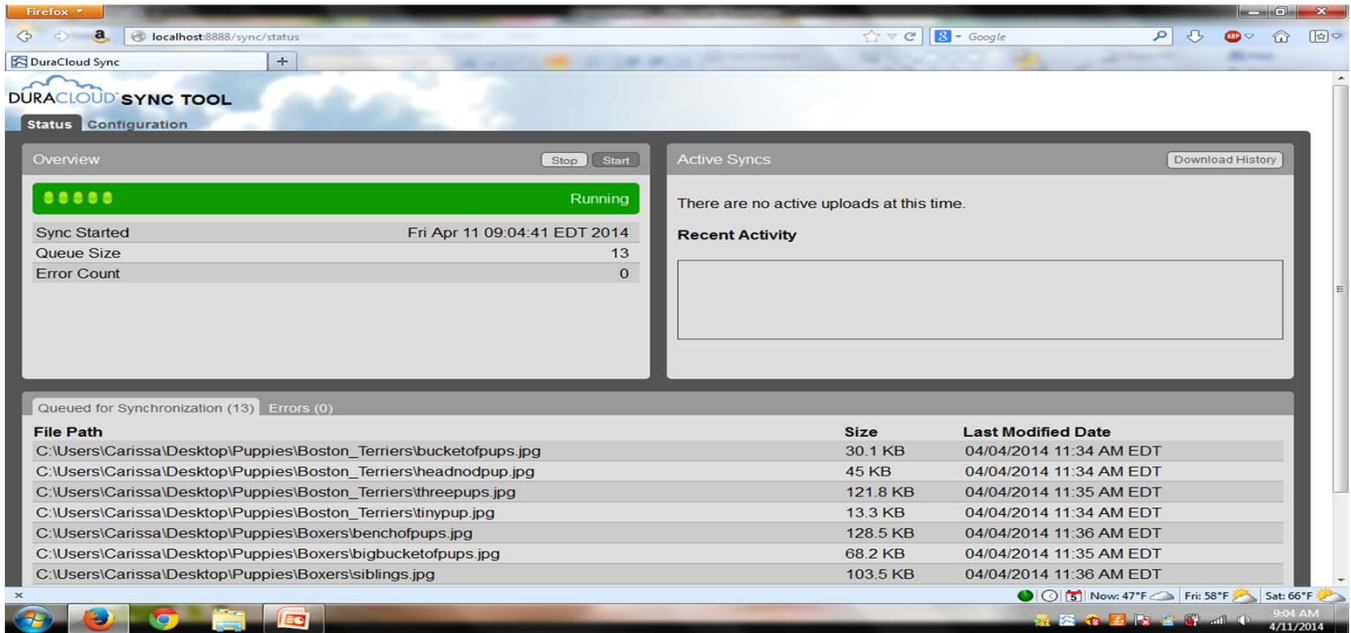
- Open Pricing
- Free Trial
- Lots of webinars and tutorials
- Learn more about the new DSpace Direct... a hosted version of the DSpace Institutional Repository software that is integrated with DuraCloud for preservation

<p><b>DuraCloud Preservation</b></p> <p>The DuraCloud Preservation plan is ideal for institutions that wish to store one copy of their content in the cloud. <i>Subscription plan is available with storage between 1-5TB of content.</i></p> <p>Example use case: <a href="#">Back-up preservation storage for a small amount of content</a></p>	<ul style="list-style-type: none"> <li>◦ <a href="#">Standard features</a></li> </ul>	<p>(Storage in Amazon S3):</p> <ul style="list-style-type: none"> <li>• \$1,875 (subscription which includes 1TB storage)</li> <li>• \$700 for additional TBs</li> </ul>
<p><b>DuraCloud Preservation Plus</b></p> <p>The DuraCloud Preservation Plus plan is best suited for organizations that wish to store two copies of their content in the cloud. <i>Subscription plan is available with storage between 1-5TB of content.</i></p> <p>Example use case: <a href="#">Archival storage for a moderate amount of content</a></p>	<ul style="list-style-type: none"> <li>◦ <a href="#">Standard features</a></li> <li>◦ <a href="#">Automatic synchronization of content between primary and secondary storage providers</a></li> <li>◦ <a href="#">Choice of secondary cloud storage providers</a></li> <li>◦ <a href="#">Automatic file recovery between copies</a></li> </ul>	<p>(Storage in Amazon S3 + Amazon Glacier):</p> <ul style="list-style-type: none"> <li>• \$2,000 (subscription which includes 1TB storage)</li> <li>• \$825 for additional TBs</li> </ul> <p>(Storage in Amazon S3 + SDSC):</p> <ul style="list-style-type: none"> <li>• \$2,875 (subscription which includes 1TB storage)</li> <li>• \$1,400 for additional TBs</li> </ul>
<p><b>DuraCloud Enterprise</b></p> <p>The DuraCloud Enterprise plan is designed to meet the needs of institutions that wish to store one copy of their content in the cloud and need to provide a variety of individuals, departments, research groups, etc. access to a single DuraCloud account. <i>Subscription plan is available with unlimited storage. Contact us for custom quote for storage beyond 10TB.</i></p> <p>Example use case: <a href="#">Long-term access storage for a variety of institutional content</a></p>	<ul style="list-style-type: none"> <li>◦ <a href="#">Standard features</a></li> <li>◦ <a href="#">Media serving</a></li> <li>◦ <a href="#">Account management</a></li> <li>◦ <a href="#">Sub-account creation</a></li> <li>◦ <a href="#">Permissions and access controls</a></li> <li>◦ <a href="#">User management</a></li> <li>◦ <a href="#">Coming Soon: Shibboleth authentication -- available to Internet2 and InCommon members</a></li> </ul>	<p>(Storage in Amazon S3):</p> <ul style="list-style-type: none"> <li>• \$5,750 (subscription which includes 1TB storage)</li> <li>• \$500 for additional TBs</li> </ul>
<p><b>DuraCloud Enterprise Plus</b></p> <p>The DuraCloud Enterprise Plus plan is intended to assist organizations that wish to store two copies of their content in the cloud and need to provide a variety of individuals, departments, research groups, etc. access to</p>	<ul style="list-style-type: none"> <li>◦ <a href="#">Standard features</a></li> <li>◦ <a href="#">Automatic synchronization of content between primary and secondary storage providers</a></li> <li>◦ <a href="#">Choice of secondary cloud storage providers</a></li> </ul>	<p>(Storage in Amazon S3 + Amazon Glacier):</p> <ul style="list-style-type: none"> <li>• \$5,875 (subscription which includes 1TB storage)</li> <li>• \$625 for additional TBs</li> </ul>

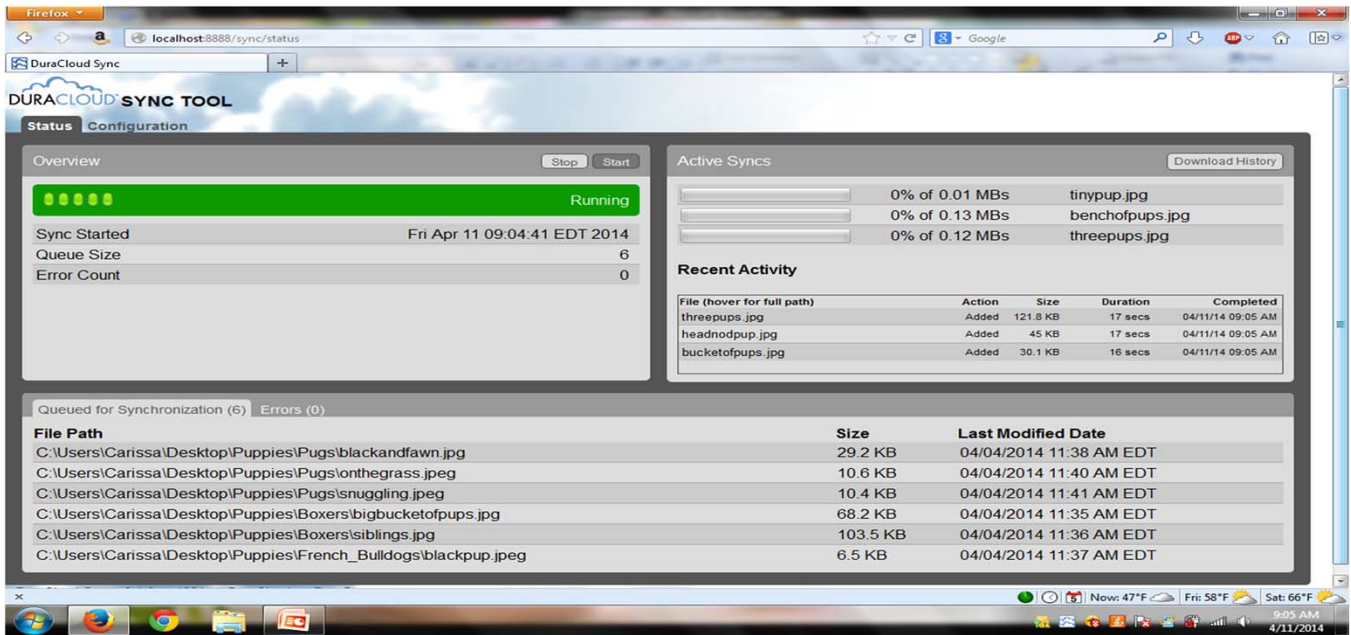
Pricing as of June 2, 2015



This screen is depicting the DuraCloud log in screen to the web interface. Every customer receives a unique log in URL (i.e. http://organizationname.duracloud.org).

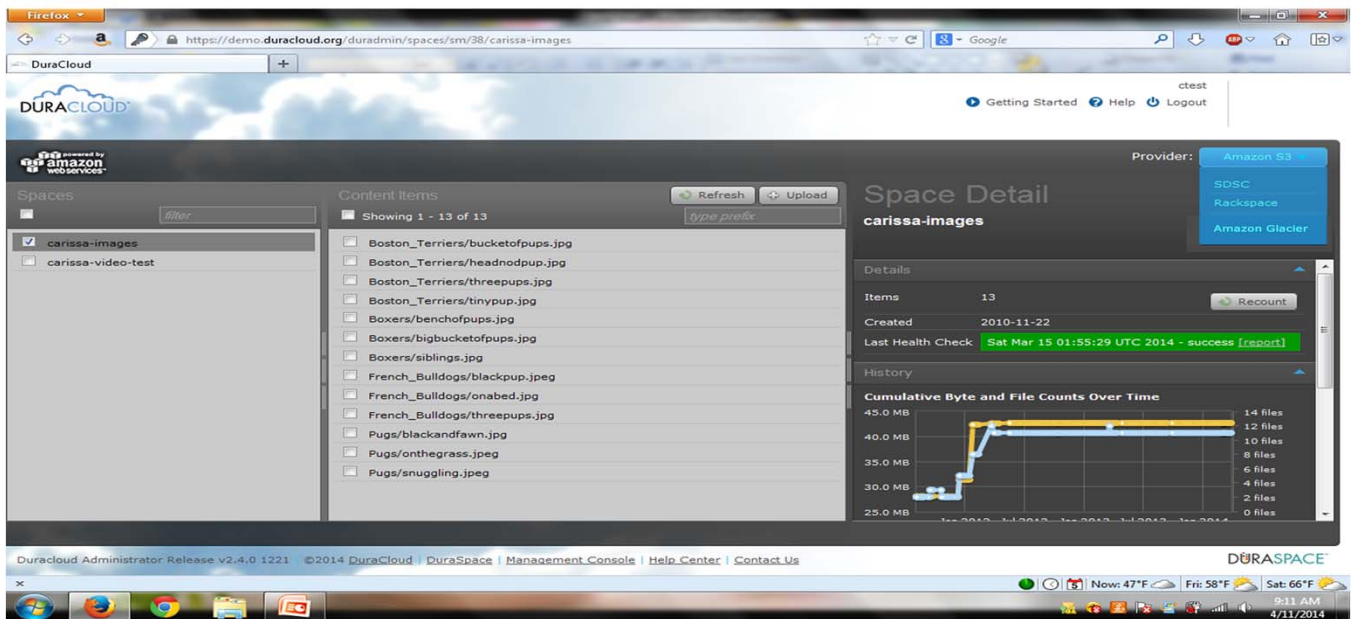


This screen is depicting the DuraCloud synchronization tool. It is a software tool that you download and install on your local machine that walks you through a “wizard” in order to select the content you wish to upload to DuraCloud. As you can see in this screen, a directory called “Puppies” is actively being uploaded to DuraCloud.

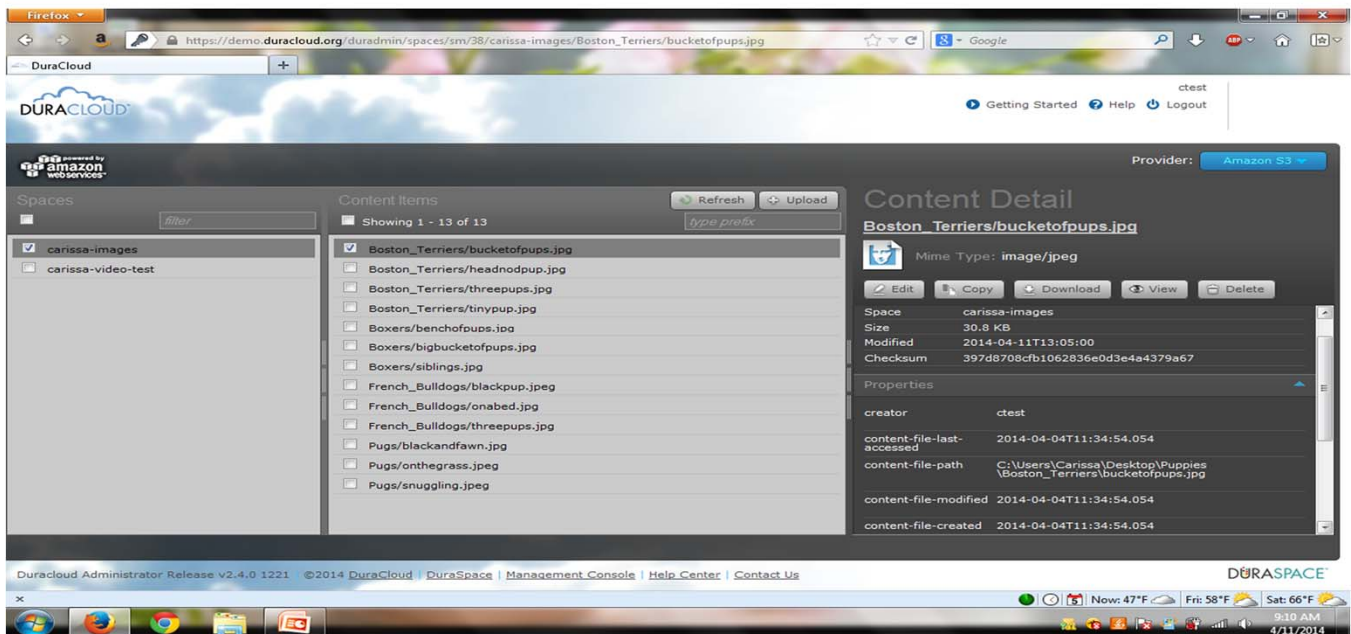


This screen is depicting the DuraCloud synchronization tool in the process of uploading content to DuraCloud. In the right hand side of the screen you can see the individual files that are actively being transferred “synced” to DuraCloud, the most recent uploads are listed below that, and the remaining files left to upload in the “queue” are listed on the bottom of the screen.





This screen is depicting the DuraCloud web interface. On the left is the “Spaces” column (spaces are DuraCloud storage containers). In the center is the “Content Items” column that lists the content stored in the space that is selected. On the right is the “Detail” column that will show the detailed information for what you’ve selected. In this case it shows the details for the “carissa-images” space that is selected (highlighted in dark gray and checked). Space details captured include: total number of items stored, data the space was created, date the last health check was run over all content (as well as a link to download the health report), and history of content additions to the space over time. Additionally, this screen depicts the various storage providers that are available with this DuraCloud account. You can navigate to the various storage providers by simply hovering over the blue provider box and clicking on the storage provider name. The DuraCloud interface will update to show the content stored in the chosen storage provider. Note that the DuraCloud interface for all providers looks/functions exactly the same. The options for cloud storage providers within DuraCloud are: Amazon S3, San Diego Supercomputer Center cloud storage, Rackspace Cloudfiles, and Amazon Glacier.



This screen is depicting the DuraCloud web interface. In this screen, the details for the “Boston\_Terriers/bucketofpups.jpg” image that is selected (highlighted in dark gray and checked) are presented in the details column on the right. Content item details captured include: space where the item is stored, the item’s size, the item’s MD5 checksum value that DuraCloud calculated upon initial upload, the user who created/uploaded the content item, and the MAC time information for the item that was provided by the local file system upon first upload (MAC = modified, access, and change times for the file). Additionally, users can edit the mime type associated with the file, create a copy of the content item to store elsewhere in DuraCloud, download the content item, view the content item within the browser (if it’s an image file type), or delete the content item from DuraCloud entirely (this is a multi-verification step process).

# Back-end/Preservation: MetaArchive

Digital POWRR Tool Evaluation Grid	Ingest			Processing						Access		Storage				Maintenance			Other				
	Copy	Fixity Check	Virus Scan	File Dedupe	Auto Unique ID	Auto Metadata Creation	Auto Metadata Harvest	Manual Metadata	Rights Management	Package Metadata	Auto SIP Creation	Public Interface	Auto DIP Creation	Auto AIP Creation	Reliable, Long-Term Bit Preservation	Redundancy	Geographically Dispersed Data Storage Model	Exit Strategy	Migration	Monitoring	Auto Recovery	Open Source	Clear Documentation
**MetaArchive (A private LOCKSS Network)	X	X		X	X		X		X	X			X	X	X	X	X		X	X		X	Varies

## Back-end/Preservation: MetaArchive

- Nonprofit; Open Pricing
- Instant community in the Cooperative!
  - All the cool kids are doing it!
- Helpful and responsive customer service
- Private LOCKSS network
- Dark Archive
- Requires dedicated IT administration
- Most memberships require attending meetings
- Assumes pre-processing work is done
- Rules for minimum processing requirements (ie file naming conventions)

What we tested

### Sample costs

<http://www.metaarchive.org/costs>

If an institution wanted to preserve 2 TB of content with us, they would pay:

**Sustaining Member:** [\$5,500 (membership) + \$2,000 (space) x 3 years] + \$4,600 (server) = \$27,100/3 years, or \$9,033/year

**Preservation Member:** [\$3,000 (membership) + \$2,000 (space) x 3 years] + \$4,600 (server) = \$19,600/3 years, or \$6,533/year

**Collaborative Member (mock example comprised of 20 institutions):** [\$4,000 (membership) + \$2,000 (space) x 3 years] + \$4,600 (server) = \$22,600/3 years or \$7,533/year total or \$377.00/year per member

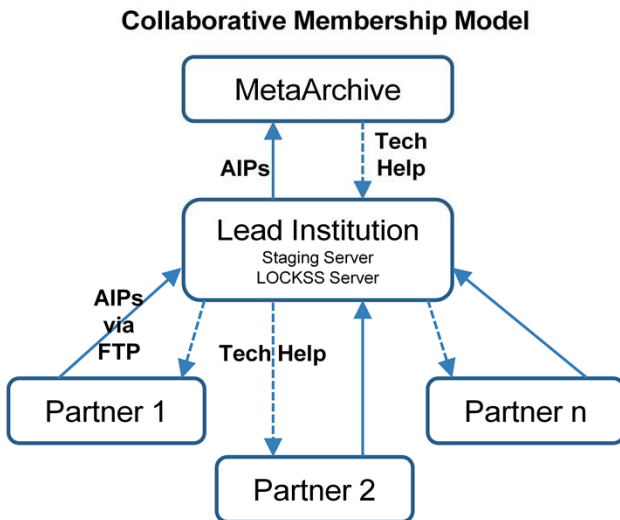
Bagger + bad file name = failed on ingest to MA

We are creating the legal framework and business model for institutions wanting to band together to join as a Collaborative Membership!!

This is the model we tested...there are other options out there for membership (standalone memberships) screen shot of pricing One point about what we tested (collaborative) might mention that standalone and PLN versions require commitment to becoming host/server for a node.

# Back-end/Preservation: MetaArchive

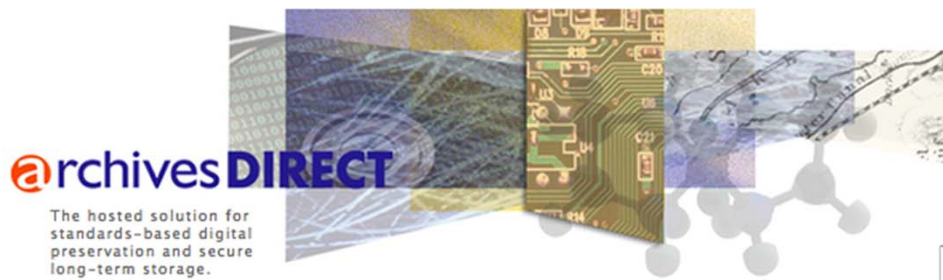
## Very simplified version of how it works:



- 1) Partners prepare their content for preservation and package it.  
→ We used the BagIt specification, and Bagger helped us with this
- 2) Lead Institution prepares a staging server, sets appropriate access protocols and assists Partners with technical help.
- 3) Partners FTP their AIP's (Bags) to the staging server at the Lead Institution.  
→ We used Filezilla
- 4) MetaArchive harvests the AIP's from the Lead Institution's staging server and pushes it into their LOCKSS network.

*One other thing:* The Lead Institution also has a dedicated server that runs the LOCKKS software, is hooked into the MetaArchive network of servers across the globe, and is actively preserving the content of other Members.

**Archivematica (Front end)**  
**+**  
**DuraCloud (Back end)**  
**=**



Archivematica and DuraCloud have teamed up to provide a hosted (web-based), “soup-to-nuts” solution that combines the front-end processing of Archivematica and the back-end preservation services of DuraCloud. This new solution is new as of February 2015, and so was not part of our original testing period.

## SUBSCRIPTION PLANS

The following subscription plans are available for the ArchivesDirect service. If you wish to receive a quote or have any questions about the ArchivesDirect subscription plan options, features, or prices, please [contact us](#).

In addition to the subscription plans below, there are optional ArchivesDirect [add-on packages](#) available.

[Subscribe](#)

Subscription Plan	Features	Annual Price
<b><a href="#">ArchivesDirect Standard</a></b> The ArchivesDirect standard plan is ideal for institutions with diverse digitized and born-digital holdings, including images, text files, office documents, PDF files, audio and video files, and forensic disk images. Users of this service will have access to a robust suite of digital preservation functions via a hosted instance of Archivemata. Archivemata is well known for its ability to produce highly standardized and interoperable Archival Information Packages (AIPs). AIP storage will be DuraCloud with secure, replicated storage in Amazon S3 and Amazon Glacier.	<a href="#">One Annual Hosted Archivemata Instance</a> <a href="#">Annual Storage: 1 TB</a> <a href="#">One Training Session</a> <a href="#">Six Hours of Customized Training and Consulting</a>	\$11,900
<b><a href="#">ArchivesDirect Digital Preservation Assessment</a></b> This plan is ideal for institutions just starting out with digital preservation or considering multiple preservation solutions. Use this service as an opportunity to learn more about digital preservation using Archivemata, to assess your readiness for digital preservation, and to plan your preservation workflows. Finish the three-month period with institution-specific use cases, workflow plans and sample Archival Information Packages (AIPs).	<a href="#">One Three-Month Hosted Archivemata Instance</a> <a href="#">One Training Session</a> <a href="#">Ten Hours of Customized Training and Consulting</a> <a href="#">Three-Month Storage: 250 GB</a>	\$4,500

Subscription plans available. They offer an add-on package for additional storage. For more than 1 TB of data, storage starts at \$1000/TB/year (this is in addition to the standard plan). Ideal for institutions with large collections. For institutions with 10 TB or more, the price of storage can be reduced further (must contact them for quote).

Pricing current **as of June 2, 2015**

\*\*If there is time, we could watch the 3-minute overview video:  
<https://www.youtube.com/watch?v=u7Ryyo2UWGA&feature=youtu.be>





# Front-end & Back-end: Preservica

- All encompassing:
  - Ingest
  - Processing
  - End-User Access
  - Preservation
  - Migration
- Aligned with OAIS reference model
- Hosted Service (Requires little IT support on your end)
- Very user friendly
- Intuitive workflows
- Exit strategy available (batch export)
- Ability to harvest via web crawls
- Solid customer support
- Different training options available for institutions with smaller budgets
- Currently uses only Amazon cloud storage
  - new options forthcoming
- Proprietary, vendor-based
- E-mail Preservation

Do we want to note that this is not a replacement for an institutional repository? Or is it?

\*\* Get a blurb on their e-mail preservation \*\*

It can also ingest, process, migrate, and preserve e-mails, including all attachments and in-line objects

**From the website:** “Preservica Cloud Edition is an affordable, secure, and easy-to-use OAIS Digital Preservation and Public Access solution hosted for you in the Cloud. You get everything you need right out-of-the-box allowing you to get started without needing to purchase in-house servers or have local IT resources.”

Starter	up to 250 GB includes training, support & user group	\$3,950 per year in Amazon S3
Starter Plus	up to 500 GB includes training, support & user group	\$6,950 per year in Amazon S3
Standard	1 to 10 TB includes training, support & user group	\$11,950 per year for first TB in Amazon S3 then \$1,450 per TB S3 per year and/or \$550 per TB Glacier per year
Volume	10 TB+ includes training, support & user group	\$ Discounted Plans (up to 50%) in Amazon S3 and/or Amazon Glacier

#### All plans include the following:

- ✓ One day in depth training on our user webinars
- ✓ Active user group with regular meetings and community portal
- ✓ Unlimited support by email
- ✓ Professional software maintenance and regular new feature upgrades
- ✓ Fully integrated suite of easy-to-use OAIS conforming workflows
- ✓ Public access/discovery module
- ✓ Active Preservation technology
- ✓ Automated DSpace ingest
- ✓ Automated CONTENTdm ingest
- ✓ Automated Microsoft SharePoint ingest
- ✓ Automated Microsoft Outlook ingest
- ✓ Advanced Website Harvesting
- ✓ CALM catalogue synchronization
- ✓ Large file upload agent
- ✓ Linked Data Registries
- ✓ All hosting, backup and operations
- ✓ Metadata held in fast Amazon RDS storage
- ✓ Durable Cloud storage - multiple copies, multiple locations, integrity checking
- ✓ Choice of Amazon S3 and/or Amazon Glacier (above 1 TB) for optimal storage costs
- ✓ Copy back to local server option
- ✓ No penalty upgrades to higher plans
- ✓ No cost to retrieve content if you decide to leave the service

Pricing and features **current as of June 2, 2015**. This cloud edition is best for small to mid-size institutions. They have options for larger institutions. New features include: Automated CONTENTdm, Microsoft SharePoint, and Outlook ingests (e-mail preservation!); large file upload; one-day introductory training on regular webinars. They have recently changed their storage (used to be starter was up to 100 GB – now 250 GB)



# INGEST

Start   Waiting   Running   **Completed**   Reports   Manage

## Workflow Details

<b>Workflow Context</b>	Discovery Package Ingest ...
<b>Workflow Definition</b>	Ingest Workflow (Manual Selection)
<b>Workflow ID</b>	474
<b>Workflow State</b>	Completed
<b>Date Started</b>	05.02.14 10:30:19
<b>Date Finished</b>	10.03.14 15:43:32
<b>Number of Files</b>	31
<b>Total Size</b>	14 MB
<b>Collection Code</b>	RC278
<b>Submission name</b>	RC 278 Kishwaukee Kiwanis
<b>Top Level Record</b>	RC 278 Kishwaukee Kiwanis

[Back](#)

## Step Progress

State	Name	Progress	Started	Finished	Messages
-------	------	----------	---------	----------	----------

<b>Collection Code</b>	RC 315
<b>Submission name</b>	RC 315 League of Women Voters of the Rochelle Area
<b>Top Level Record</b>	RC 315 League of Women Voters of the Rochelle Area

[Back](#)

### Step Progress

State	Name	Progress	Started	Finished	Messages
✓	Import from Transfer Area	<div style="width: 100%; height: 10px; background-color: green;"></div>	07.03.14 14:33:09	07.03.14 14:33:14	
✓	Virus Check	<div style="width: 100%; height: 10px; background-color: green;"></div>	07.03.14 14:33:14	07.03.14 14:33:47	
✓	Fixity Check	<div style="width: 100%; height: 10px; background-color: green;"></div>	07.03.14 14:33:47	07.03.14 14:33:50	
✓	Metadata Integrity	<div style="width: 100%; height: 10px; background-color: green;"></div>	07.03.14 14:33:50	07.03.14 14:33:53	
✓	Content Integrity	<div style="width: 100%; height: 10px; background-color: green;"></div>	07.03.14 14:33:53	07.03.14 14:33:56	
✓	SIP Validation	<div style="width: 100%; height: 10px; background-color: green;"></div>	07.03.14 14:33:56	07.03.14 14:33:59	
✓	SIP Validation with Database Crosscheck	<div style="width: 100%; height: 10px; background-color: green;"></div>	07.03.14 14:33:59	07.03.14 14:34:02	
✓	Characterise	<div style="width: 100%; height: 10px; background-color: green;"></div>	07.03.14 14:34:02	07.03.14 14:34:17	<a href="#">View</a>
✗	Store Files	<div style="width: 100%; height: 10px; background-color: red;"></div>	07.03.14 14:34:17	08.03.14 05:09:08	<a href="#">View</a>

If there is an error during the Ingest, the system will notify you and show an error message. You can view messages on the far right column.

### Preservation Plan Details

<b>Transformation Name</b>	9605ea1e-2959-4c0f-b6a1-73fd0b6aad97
<b>Description</b>	Test1
<b>Date Created</b>	
<b>Number Of File Sets</b>	1
<b>Migration Pathway Role</b>	Preservation
<b>Usage</b>	Test
<b>Top-Level Collection</b>	Danielle

This transformation is ready for execution.

#### Selected File Sets

	Collection Title	Catalogue Ref	Description	Manifestation	Number of Files	Files at Risk	Size	Size of Files at Risk	Details
	Danielle	RC 278 Kishwaukee Kiwanis	RC 278 Kishwaukee Kiwanis	1	31	2	15.41 MB	592.03 kB	<input type="button" value="Details"/>
Totals	1				31	2	15.41 MB	592.03 kB	

#### Selected Migration Pathways

Source Format	Version	PUID	Target Format	Version	PUID	Tool
Microsoft Word for Windows	2007 onwards	fmt/412	Acrobat PDF 1.4 - Portable Document Format	1.4	fmt/18	Open Office PDF

Preservation Plan details – deciding what to put in the AIP.

- Identify what is most at risk
- Migrate files to a preservation-friendly format.







Start    Waiting    **Running**    Completed    Reports    Manage

### Workflow Details

<b>Workflow Context</b>	Migration (Filtered)
<b>Workflow Definition</b>	Preservation Workflow (Filter by Top-Level Collection)
<b>Workflow ID</b>	879
<b>Workflow State</b>	Active
<b>Date Started</b>	13.04.14 15:47:58
<b>Date Finished</b>	
<b>Number Of File Sets</b>	1
<b>Migration Pathway Role</b>	Preservation

### Step Progress

State	Name	Progress	Started	Finished	Messages
	New Details	<div style="width: 100%; height: 10px; background-color: green;"></div>	13.04.14 15:47:58	13.04.14 15:53:39	
	Pick Formats at Risk	<div style="width: 100%; height: 10px; background-color: green;"></div>	13.04.14 15:53:39	13.04.14 15:53:49	
	Pick Formats at Risk	<div style="width: 100%; height: 10px; background-color: green;"></div>	13.04.14 15:53:49	13.04.14 15:55:52	
	Pick File Sets	<div style="width: 100%; height: 10px; background-color: green;"></div>	13.04.14 15:55:52	13.04.14 15:58:50	
	Ready	<div style="width: 100%; height: 10px; background-color: green;"></div>	13.04.14 15:58:50	13.04.14 15:59:58	
	Migrate AIPs	<div style="width: 50%; height: 10px; background-color: gray;"></div>	13.04.14 15:59:59		

PRESERVATION - Migrate to AIP in Progress

Back

### Step Progress

State	Name	Progress	Started	Finished	Messages
	New Details	<div style="width: 100%;"></div>	13.04.14 15:47:58	13.04.14 15:53:39	
	Pick Formats at Risk	<div style="width: 100%;"></div>	13.04.14 15:53:39	13.04.14 15:53:49	
	Pick Formats at Risk	<div style="width: 100%;"></div>	13.04.14 15:53:49	13.04.14 15:55:52	
	Pick File Sets	<div style="width: 100%;"></div>	13.04.14 15:55:52	13.04.14 15:58:50	
	Ready	<div style="width: 100%;"></div>	13.04.14 15:58:50	13.04.14 15:59:58	
	Migrate AIPs	<div style="width: 100%;"></div>	13.04.14 15:59:59	13.04.14 16:01:07	<a href="#">View</a>

### Preservation Plan Details

<b>Transformation Name</b>	9605ea1e-2959-4c0f-b6a1-73fd0b6aad97
<b>Description</b>	Test1
<b>Number Of File Sets</b>	1
<b>Migration Pathway Role</b>	Preservation
<b>Usage</b>	Test

The transformation has completed.







Top Level Record	Collection Title	Outcome
RC 278 Kishwaukee Kiwanis	Danielle	Transformation Complete

PRESERVATION - Storing to the AIP successful!





## Saved Reports

Report Name	Report Summary	Download
Deleted manifestations	Show deleted redundant manifestations and associated files	<a href="#">Parameters</a>
Download activity	Download activity summary	<a href="#">Parameters</a>
File Download activity	File Download activity summary	<a href="#">Parameters</a>
File formats	Show the breakdown of file formats in the archive	  
File formats (details)	Shows the file formats with version and puid details	  
Files At Risk	Show deliverable units containing 'at risk' files in their active manifestation	<a href="#">Parameters</a>

It provides nice reports!

Digital Preservation

Help Home English


Tools View Manifestation Preservation 1 (Active, Original)

Kish-Newslett... Kish-Newslett... Kish-Newslett... Kish-Newslett... Kish-Newslett... Kish-Newslett...


Kish-Newslett... Kish-Newslett... Kish-Newslett... Kish-Newslett... Kish-Newslett... Kish-Newslett...

Filter Apply  Show accrued files  Show children  Show deprecate 1 of 2 1-20 of 31

Name: Kish-Newsletter-2013-09-25.doc  
File Type: Microsoft Word Document  
Type: File Size: 2.359 MB Reference: 17d420e1-0910-4887-a2e8-104e0bf9a338



Access view – you can actually SEE what was migrated!! You can also migrate and download documents to create access copies.


[dorsetarchives.com](#)
[browse archive](#)


Search Our Archive

[search](#)

---

[Twitter](#)  
[Facebook](#)  
[LinkedIn](#)  
[RSS](#)

FEATURED



### New Dorset Photo Collection

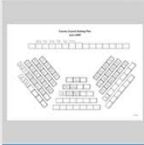
Please check out our new photo collection of donated photos from around the county supplied for free use. These photos have been provided by professional and amateur photographers and cover the period 1990-2002.

Object Type: Record  
In Collection: Dorset Archives

[details](#) [view images](#)

---


Other items of interest



#### 2009 Council Seating

Council seating and processes


Object Type: File  
In Record: Public Documents



#### Ancestry Documents

Collection of information supplied by Ancestry.com


Object Type: Record  
In Collection: Dorset Archives



#### Dorset Strategic Partnership Annual Review

Audit & Scrutiny Committee: Dorset Strategic Partnership Annual Review and Use of Funding

Object Type: File  
In Record: Public Documents



#### William Barnes Collection

William Barnes is best known for his dialect poetry, though his formal schooling ended at the age of 13. The collection at the Dorset History Centre contains most of his published works.

Object Type: Record  
In Collection: Dorset Archives

Access interface

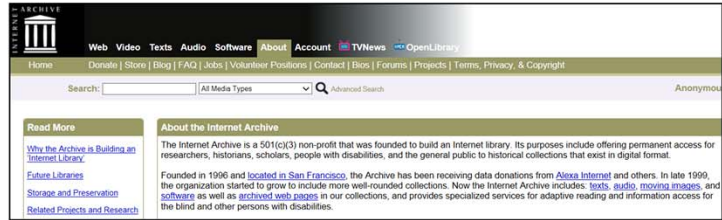
# Front-end & Back-end: Internet Archive

Digital POWRR Tool Evaluation Grid	Ingest			Processing						Access		Storage				Maintenance			Other				
	Copy	Fixity Check	Virus Scan	File Dedupe	Auto Unique ID	Auto Metadata Creation	Auto Metadata Harvest	Manual Metadata	Rights Management	Package Metadata	Auto SIP Creation	Public Interface	Auto DIP Creation	Auto AIP Creation	Reliable, Long-Term Bit Preservation	Redundancy	Geographically Dispersed Data Storage Model	Exit Strategy	Migration	Monitoring	Auto Recovery	Open Source	Clear Documentation
**Internet Archive	x	x			x	x	x	x	x	x	x	x	x	x	x	x					x		Free

This is meant for very small institutions. Internet Archive will only take public domain material.

# Internet Archive

- Only intended for materials in the public domain (available to everyone).
- Geographically distributed copies.
- No frills (and no charge!) service.



**Share your Files**

Please contribute books, audio, and video files that you have the right to share. The Internet Archive, a non-profit library, will provide free storage and access to them. (We reserve the right to remove any submitted material.)

[Upload Files](#)

**Or choose an alternate upload method:**

<a href="#">Classic Uploader</a>	Internet Explorer users and print disabled (prefer to <a href="#">not use</a> )
<a href="#">Live Music Archive</a>	For etree items
<a href="#">FTP</a>	Large file alternative

- Handles books best, but can accommodate manuscripts, audio, video, and images.
- Is especially suited for small (VERY small institutions with limited (or no) other alternatives.
- Does offer a more robust preservation product through its Archive-It service.

Stores a preservation copy of the uploaded materials on 2 different continents  
We have a tutorial for how to do this !!!!

Account TVNews OpenLibrary  
 Forgot Password | Archive Bookmarks  
 Advanced Search Hello Northern Illinois University Library (not you? sign out)

<b>Page Title *</b>	postelpopulism1	✓	Drag and Drop More Files Here or <a href="#">Select files to add</a> <table border="1"> <thead> <tr> <th>Name</th> <th>Size</th> <th>x</th> </tr> </thead> <tbody> <tr> <td>postelpopulism1.avi</td> <td>1.9 GB</td> <td>✗</td> </tr> </tbody> </table>	Name	Size	x	postelpopulism1.avi	1.9 GB	✗
Name	Size	x							
postelpopulism1.avi	1.9 GB	✗							
<b>Page URL *</b>	http://archive.org/details/postelpopulism1	✓							
<b>Description *</b>	Add a description of the item page	✓							
<b>Subject Tags *</b>	Add keywords, separated by commas	✓							
<b>Creator</b>	Creator of the content	✓							
<b>Date</b>	Date work was created/published	✓							
<b>Collection *</b>	Community Video	✓							
<b>Test Item</b>	No	✓							
<b>Language</b>	Language of the work	✓							
<b>License</b>	No license selected	✓							
<b>More Options</b>	<a href="#">Add additional metadata...</a>								

[Upload and Create Your Item](#)

[Please provide feedback about the new Beta Uploader](#) — [Instructions on how to preset metadata](#) — [Save this metadata](#)

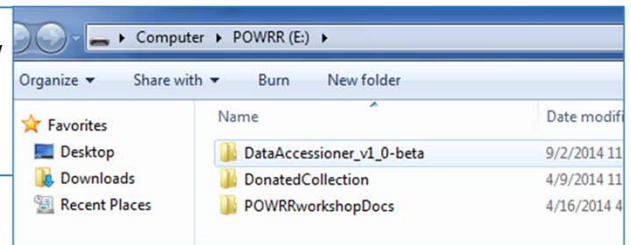
**Mention they have a free version**

**LUNCH**

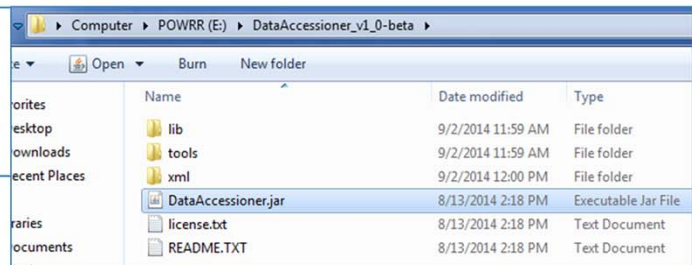
# Activity: Accessioning a Digital Collection 1 - 2pm

## 1. Insert flash drive and open the explorer window

Data Accessioner  
Donated Collection Folder  
Digital Collections Inventory file  
Other stuff.....



## 2. Navigate to DataAccessioner.jar and open it



**Take out the printed handout so you can follow along!**

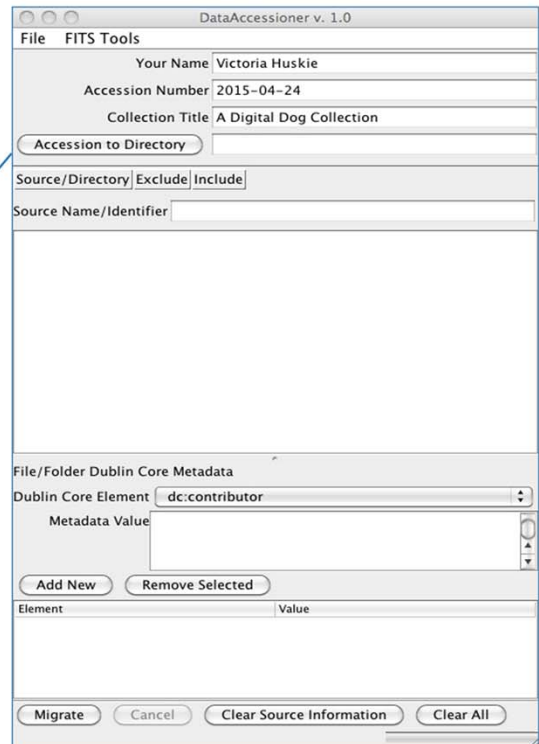
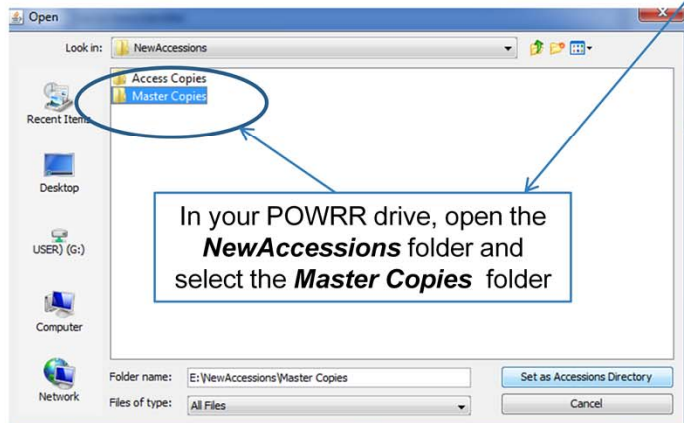
Walk through opening DataAccessioner → DataAccessioner.jar

\*\*The morning session we accessioned a collection of cats. This afternoon is for you dog lovers. 😊\*\*



# Create your accession directory: Where you want the collection to go live

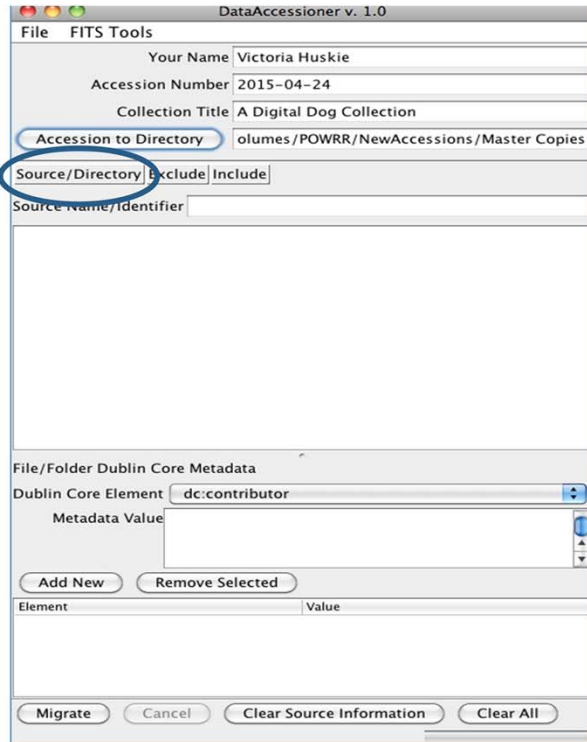
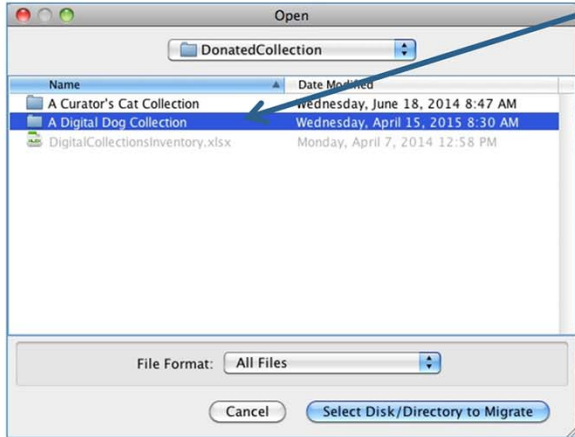
*Preferably a stable media like your network drive*



Walk through opening DataAccessioner → DataAccessioner.jar

Name this folder NewAccessions. Also make an Access Copies and Master Copies Folder.

# Select the collection you are accessioning



## Populate descriptive metadata and migrate your collection

Select which element you want to add metadata to

Add the Dublin Core Metadata goes here

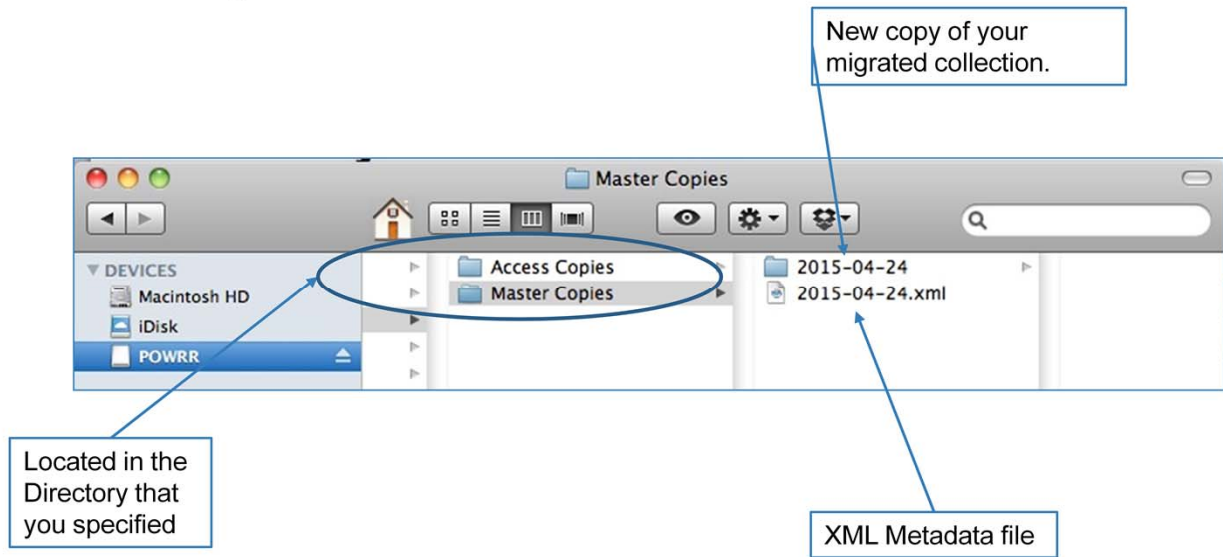
Hit the "Migrate" button to begin the migration process.

Element	Value
dc:creator	Jane and John Moneybags
dc:date	2015
dc:description	A collection of dog images collected by J...

You will be able to see the progress bar move at the bottom.

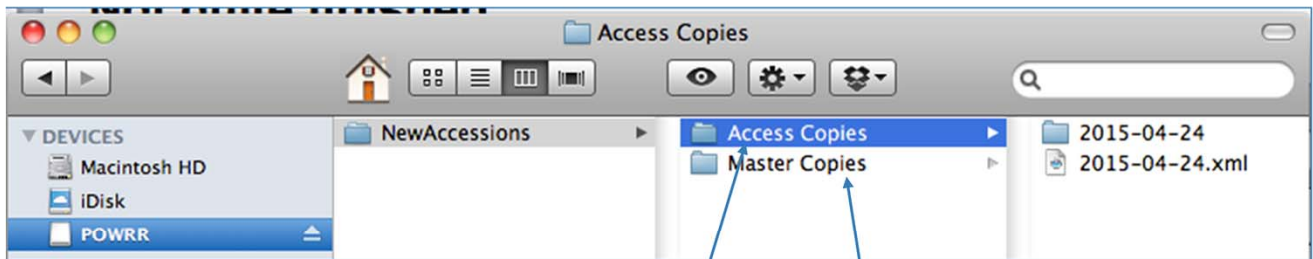
Mention you can copy this in from the "basic" accession file created earlier.

# What did you create?



**You. Are. AWESOME.**

## Not quite finished...



Make a copy of the Master, place in the Access Copies folder, and don't touch the Master Copy again unless a new derivative is needed or until you move it into a preservation system!!

Make a copy of the Master, place in the Access Copies folder, and don't touch the Master Copy again unless a new derivative is needed, or until you move it into a preservation system.

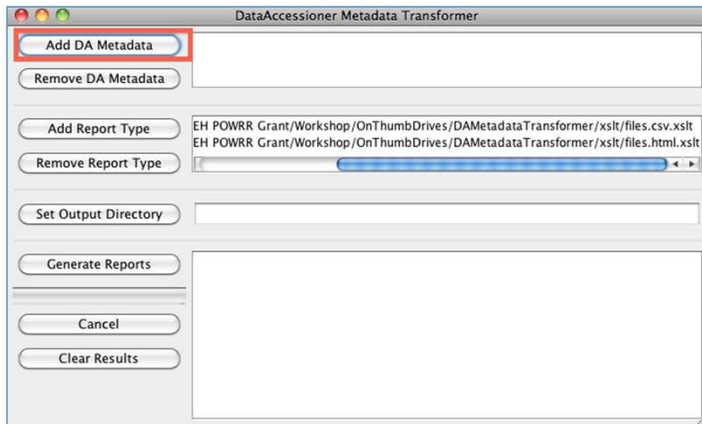
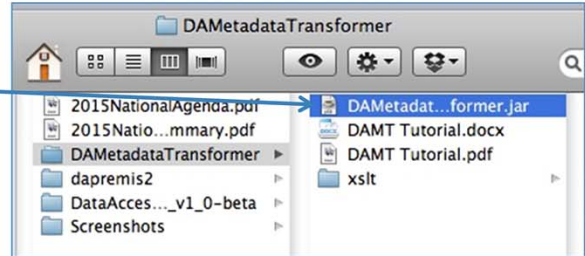
And finally...update your Inventory to reflect the location of the Access Copy. Note addition of XML file after processing.

Category	Title & Description	Date	Location	Extent	Format
(locally defined; project name? content creation method?)	(Donor applied and/or yours... what's your local practice?)	(YYYYMMDD or other locally defined format for accession date)	(Storage place of choice -- networked server recommended)	(Quantity of folders, files, by type or total size)	(What extensions are involved: .jpg, .tif, .xls?)
Special Collections, mixed; digitized and born digital	A Digital Dog Collection. Donated by Jane (nee Pennypincher) and John Moneybags, Class of 2006. Various images and visual materials about dogs of interest to the donors; some material may have copyright restrictions by law	2015-04-24	C:\Users\User\Desktop\NewAccession\Masters	12.93 MB	19 jpgs; 1 ogg; 1 png; 1 xml file from DA
Special Collections, mixed; digitized and born digital	A Digital Dog Collection. Donated by Jane (nee Pennypincher) and John Moneybags, Class of 2006. Various images and visual materials about dogs of interest to the donors; some material may have copyright restrictions by law	2015-04-24_AccessCopie	C:\Users\User\Desktop\NewAccession\AccessCopies	12.93 MB	19 jpgs; 1 ogg; 1 png; 1 xml file from DA

Add the Access Copies to your inventory spreadsheet.

# DA: Metadata Transformer

Navigate to DAMetadataTransformer.jar and open it

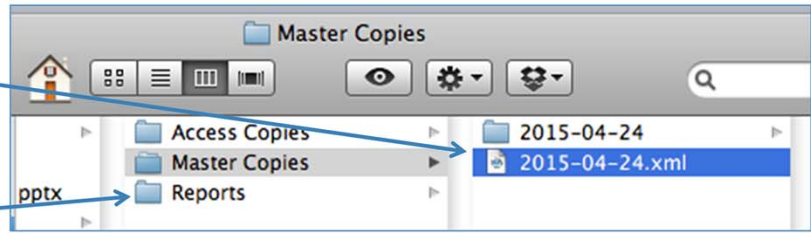


Click on "Add DA Metadata" button

Remember, DA only generates a raw XML file. For a user-friendly version to read the metadata, we will be using this reporting function called DataAccessioner: Metadata Transformer tool.

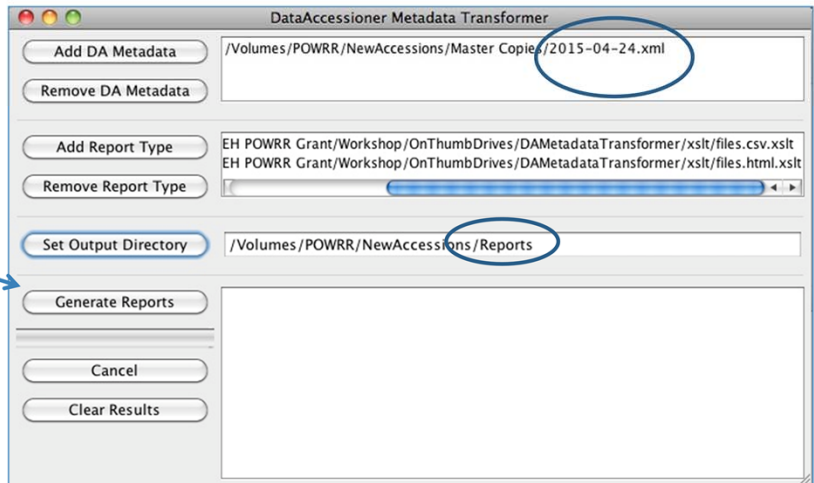
Navigate to the XML file you just created.

We also want a place where the new reports will go live. In this case, you can create a "Reports" folder



You will see the XML created and the folder where the new reports will go live.

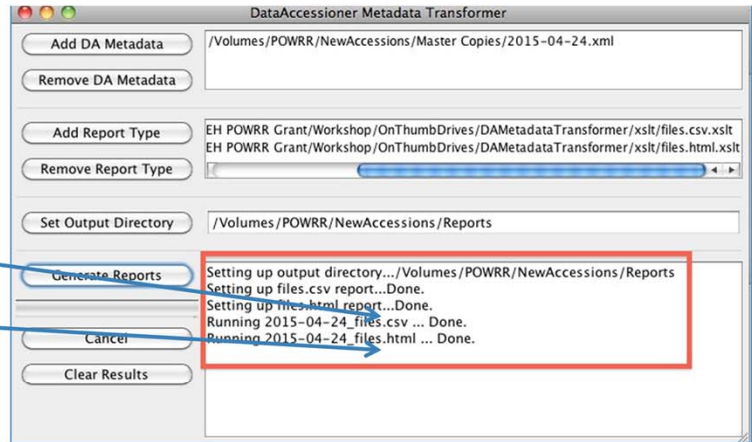
Click on "Generate Reports"



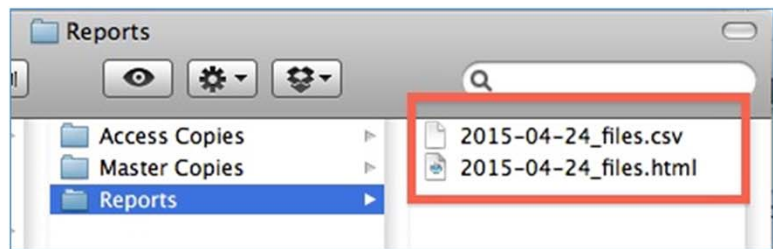


Once the reports have been generated, DA:MT will show the results.

In this case, two files were created: a CSV file and HTML file.



To view the files, navigate back to the "Reports" folder.



## CSV FILE

	A	B	C	D	E	F	G
1	directory path	file name	last modified	size (bytes)	md5	file format	
2	A Digital Dog Collection/Action Shots/	Catch_the_ball.jpg	2015-03-23T15:08:52.000	28095	e72655088fc84f685	JPEG File Interchange Format	
3	A Digital Dog Collection/Action Shots/	Dachshund_leaping_from_log.jpg	2015-03-23T15:02:21.000	1213250	8c42ff7cfa14dea90	Exchangeable Image File Format	
4	A Digital Dog Collection/Action Shots/	Play_time.jpg	2015-03-23T14:51:31.000	1140632	fd8b12e0789c74e32	JPEG File Interchange Format	
5	A Digital Dog Collection/Action Shots/	All_Dressed_Up.jpg	2015-03-23T15:03:31.000	10466	3b87fd7a0982700e	JPEG File Interchange Format	
6	A Digital Dog Collection/Action Shots/	Animal-Loocomotion-Photo-Cat-and-dog-jumping-7	2015-03-23T14:59:42.000	142754	1d8b90918b7fd002c	JPEG File Interchange Format	
7	A Digital Dog Collection/Action Shots/	Better_call_a_doctor_whi.jpg	2015-03-23T15:22:24.000	64357	9e53d0749f7cbd60e	JPEG File Interchange Format	
8	A Digital Dog Collection/Action Shots/	Blizzards_can_be_fun_whi.jpg	2015-03-23T15:24:20.000	15265	507ae4998624c05f5	JPEG File Interchange Format	
9	A Digital Dog Collection/Action Shots/	Boxer_dog_footage_1901.ogg.160p.ogv	2015-03-23T14:45:13.000	1864745	78e5d16b897570e3	Ogg data, Skeleton v4.0	
10	A Digital Dog Collection/Action Shots/	Historic_Dogs.png	2015-03-23T15:05:57.000	1044189	e02c655f61554a09e	Portable Network Graphics	
11	A Digital Dog Collection/Action Shots/	Small_dog_on_a_pedestal.jpg	2015-03-23T15:26:56.000	16305	017ccfd735f58304d	JPEG File Interchange Format	
12	A Digital Dog Collection/Action Shots/	Tootsie_takes_up_music_whi.jpg	2015-03-23T15:23:35.000	61558	c6729ecbf3f97941d	JPEG File Interchange Format	
13	A Digital Dog Collection/Action Shots/	A-Corgie-floating-along_(2509104391).jpg	2015-03-23T14:53:03.000	154062	c8af18afc2693e59b	Exchangeable Image File Format	
14	A Digital Dog Collection/Action Shots/	DP_doge.jpg	2015-03-23T15:47:59.000	160810	8b3bd9afad0cfa723	JPEG File Interchange Format	
15	A Digital Dog Collection/Action Shots/	Husky_Yawn.jpg	2015-03-23T14:58:56.000	572373	85562ffa27e2d6754	JPEG File Interchange Format	
16	A Digital Dog Collection/Action Shots/	Marley.jpg	2015-03-23T14:57:05.000	48366	43d2a524ea4f42fd8	JPEG File Interchange Format	
17	A Digital Dog Collection/Action Shots/	Mission_NIU.jpg	2015-03-23T15:29:31.000	117662	6f8bd3a9cbd7a1055	JPEG File Interchange Format	
18	A Digital Dog Collection/Action Shots/	Twinsies.jpg	2015-03-23T15:07:46.000	37563	9d9d8ab31c1c2675	JPEG File Interchange Format	
19	A Digital Dog Collection/Action Shots/	Watermelon_thief.jpg	2015-03-23T15:07:59.000	33997	b79becf2b6972c079	JPEG File Interchange Format	
20	A Digital Dog Collection/Action Shots/	Bath_time.jpg	2015-03-23T15:10:52.000	28354	6c939fe1abeffb9a81	JPEG File Interchange Format	
21	A Digital Dog Collection/Action Shots/	Brown_puppy_(9899551176).jpg	2015-03-23T14:48:59.000	6134224	c36586d1f2f3b670e	Exchangeable Image File Format	
22	A Digital Dog Collection/Action Shots/	Young_Marley.jpg	2015-03-24T08:52:45.000	34482	3a66d55a12236fe37	JPEG File Interchange Format	

Much prettier than the raw XML, right?

**\*\*GO TO A LIVE VERSION TO DEMO THE HTML FILE. ONLY WORKS WITH CHROME\*\***

# **CONGRATULATIONS!**

**We call this “Digital Preservation in Your Office”**

**There are things that need to happen *outside* of  
your office as well....**

**Short break? Assess the room and the time to see if people want a stretch break.**

# Outside Your Office

Digital Preservation is not sustainable by just using a tool or selecting a service. Sustainability takes funding and people.

You cannot do this alone. **You will need to talk to other people...** because you are not the only boss of this.

Successful Digital Preservation programs take a team of people at multiple administrative levels.

Three-Legged Stool of Digital Preservation



Anne R. Kenney  
Nancy McGovern

Digital Preservation Management Workshop  
<http://www.dpworkshop.org/>

A fully implemented and viable preservation program addresses organizational issues, technological concerns, and funding questions, balancing them like a three-legged stool. **Organizational Infrastructure** includes the policies, procedures, practices, people—the elements that any programmatic area needs to thrive, but specialized to address digital preservation requirements. It addresses this key development question:

**What** are the requirements and parameters for the organization's digital preservation program?

**0101 Technological Infrastructure** consists of the requisite equipment, software, hardware, a secure environment, and skills to establish and maintain the digital preservation program. It anticipates and responds wisely to changing technology. It addresses this key development question:

**How** will the organization meet defined digital preservation requirements?

**\$\$\$\$ Resources Framework** addresses the requisite startup, ongoing, and contingency funding to enable and sustain the digital preservation program. It addresses this key development question:

**What** resources will it take to develop and maintain the organization's digital preservation program?

# Assemble Your Team!



Image: Flickr Commons

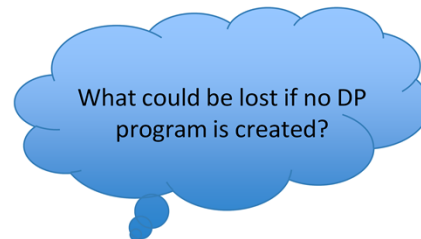
Know who your allies are and how their role can help you towards your DP goals. Because you wouldn't want to send Aqua Man into a land-based battle.

# Outside Your Office

## *Group Activity: 3-3-3 Action Plan*

1.) Create a list of all roles in an organization that should play a part in some aspect of digital preservation

2.) Make a list of potential consequences if no action is taken



Give them a couple of minutes to brainstorm some on their own (or in small groups based on institution type...)

Start white-boarding the roles – **Activity: Identify risks of inaction**

**If you don't do DP, what would happen?**

**1.) Get in groups for 5 minutes, create a list of roles. White board for 5 minutes**

**2.) IN THE SAME GROUPS, make lists of inaction. What are the risks of doing nothing? Put on white board**

### 3-3-3 Action Plan: Build Your Team

Now let's move from roles to people....

- List 3 individuals *at your institution*
- Which of these folks are you willing to contact in the next 2 weeks?
  - ...in the following month?
  - ...in the following 3 months?
- List 3 activities to accomplish in the first 3 months

I can imagine that some people might not “already have a working relationship” with people in the roles listed. Maybe add this as a point for emphasis.

Think of it as moving even one square to the right on the NDSA Levels of Preservation grid

- After bringing these colleagues on board, what are 3 concrete, small steps that you can take together to move your burgeoning DP program forward?

# Examples of Initial DP Activities

- **Conversations/Meetings**
- **Inventory what you already have**
- **Enhance the metadata of the records you already have**
- **Update current policies to include digital materials**
- **Download DA and play with it some more!**
- **Tool investigation**
- **Look at other institutions' digital preservation policies**
- **Outreach/Education at your institution**
- **Read the POWRR white paper**

- Conversations/Meetings
- *Inventory what you already have*
- Enhance the metadata of the records you already have
- *Look at how current policies address digital materials (ex. collection development)*
- Download DA and play with it some more!
- Tool investigation: Dig a little deeper on tools that piqued your interest today*
- Look at other institutions' DP policies with an eye to crafting your own
- *Engage in some outreach/education activities...host a Brown Bag!*
- Read the POWRR white paper



## **Now Let's Assess...**

### **How will you know if your 3 activities succeeded?**

- Added people to team?
- Number of people newly educated?
- Number of items added to inventory?
- Number of tools investigated?
- Number of DP policies looked at?
- Revised standing policies?

**BREAK TIME!**

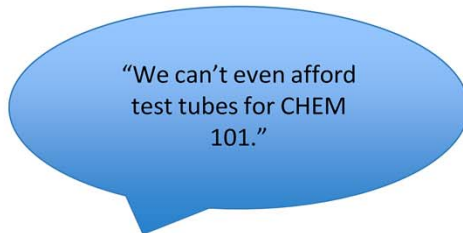
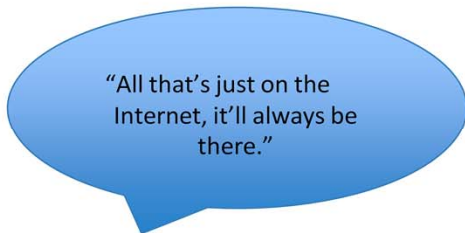
Back by 2:45, please

Next Steps:

Advocacy, Policy, and Potential Solution Models

# Next Steps: Advocacy

- Advocacy is valuable because you're educating people about why digital preservation is also THEIR problem.



\*\* These are actual conversations that came up during our attempts to do advocacy at our institutions. You may run into some roadblocks, but it's vital to show people that DP is also THEIR problem, despite all the other problems that exist.\*\*

**Have them get out one pagers here!**

Advocacy is important to teach people that it's also THEIR problem.

-Our one-pagers may help you frame why digital preservation is important to different jobs/function.

-The risks of doing nothing are a lot greater than they may think

- Good policies incorporate multiple viewpoints.
- Other people at your institutions will bring up issues – and possible solutions – you may have missed.
- You will discover many things that you don't directly control that still directly affect your work. This will lead you to more people to add to your team.

# Next Steps: Towards a Policy

**You have started assembling your team....now what?**

We found a gap analysis *really* helpful:

- Where are you now?
- Where would you ideally like to be?
- What is keeping your institution from moving in that direction?
- What are some interim steps you can take to move in the right direction?

# Next Steps: Towards a Policy

We also found that Gap Analyses can be challenging...

- Be brutally honest. It's the only way to move forward.
- Look closely at risk: What is the cost of doing nothing?
- Documenting what you know will tell you what you don't know.
- Feel free to look at our case studies and see how it worked. Our wiki has the case studies of all 5 of the POWRR partner institutions.  
[http://powrr-wiki.lib.niu.edu/index.php/Main\\_Page](http://powrr-wiki.lib.niu.edu/index.php/Main_Page)

Identifying *what* is it risk (inventory), *why* it is at risk (one-pagers), and what the loss would mean to your institution will make an incredibly compelling case to your leaders (aka the holders of the checkbook!)

# Solution in Practice is Iterative

- Not all tools and services are created equal.
- Choices of tools are *not* forever. They serve what you need now, selected with an eye to later.
- Starting small is good enough! A simple tool may still move you closer to your goals.
- Knowing what you have is crucial. Documentation more so.
- You already have many of the necessary skills!

Monologue to talk them off the edge. You don't have to pick the tools/services that will do ALL of the steps from the get go. Start with a step or two. And it doesn't even have to be with a fancy schmancy tool!

CAVEAT: Time Sensitive. Software changes quickly. Like, REALLY QUICKLY. Reference COPTR

# **Next Steps: Potential Solution Models**

JAIME

# How to Decide? Results May Vary...

## Things to consider:

- How many staff members will be actively engaged in the digital curation lifecycle? Are they tech-savvy?
- How robust and supportive is your technical/systems group? Do you even have one? How about some developers/programmers...have any of those on staff?
- Is your institution already using archival management software or an Institutional Repository (like ARCHON/ArchivesSpace, BePress, Fedora etc.)? You'll want to select tools/services that work well with what you have.
- Do you have digital collections unique to your institution that are irreplaceable? Consider organizing collections along the lines of those that warrant more robust preservation services than others. For example:

1 TB (High Value)	→	MetaArchive (gold standard)
3 TB (Medium Value)	→	Amazon Glacier (cheapest storage with fixity checking)
Rest (Replaceable)	→	Tape Drive Backups

**In other words: One tool/service will not be your only solution.**

## **This slide should be less wordy. MAKE LESS WORDY**

- How many staff members will be actively engaged in the digital curation lifecycle? Are they tech-savvy?
- How robust and supportive is your technical/systems group? Do you even have one? How about some developers/programmers...have any of those on staff?
- Is your institution already using archival management software or an Institutional Repository (like ARCHON/ArchivesSpace, BePress, Fedora etc.)? You'll want to select tools/services that work well with what you have.
- Do you have digital collections unique to your institution that are irreplaceable? Consider organizing collections along the lines of those that warrant more robust preservation services than others. For example:

1 TB (High Value)	→	MetaArchive (gold standard)
3 TB (Medium Value)	→	Amazon Glacier (cheapest storage with fixity checking)
Rest (Replaceable)	→	Tape Drive Backups



# How to Decide? Results May Vary...

Remember: Smaller institutions with less resources may also have unique advantages like....

It doesn't take years to set up an account with something like DuraCloud.

You only need to convince the person one level above you to get what you need.

Want to install a simple open source tool? Go for it!

- Less red tape for getting things done
- Fewer levels to push requests for additional resources through
- Self-administered workstations (aka no IT administrative lock downs)
- Personnel-heavy operating model (usually has smaller cash flow)
- Higher cash flows and less data (like small, private institution)

This is ideal for running a \*free\* robust tool that requires a developer and server administrator like Archivematica.

You can purchase a reasonably-priced, hosted soup-to-nuts solution.

You can have a lot of archivists and no one tech-savvy. That could be a problem

NOTE: maybe mention that Glacier is available as standalone storage and also a back-end option with X products that we tested?

## Next Steps: Potential Solution Models

**POWRR White Paper available at:**

<http://commons.lib.niu.edu/handle/10843/13610>

From Theory to Action:  
"Good Enough" Digital Preservation Solutions for  
Under-Resourced Cultural Heritage Institutions

---

A Digital POWRR White Paper for the  
Institute of Museum and Library Services  
*August 2014*

AUTHORS  
Jaime Schumacher  
Lynne M. Thomas  
Drew Vandecreek

Stacey Erdman  
Jeff Hancs  
Aaisna Haykal  
Meg Miner  
Patrice-Andre Prud'homme  
Danielle Spalenka

---



# Wrapping Up

## Our Final Thoughts & Your Questions

I survived the POWRR workshop! Now what?

<https://digitalPOWRR.niu.edu/survived-powrr-wkshp/>

We're here to help. Seriously.

YOU CAN DO THIS. Really. But not alone. So bring some friends.

*"If you want to go fast...go alone. If you want to go far...go together."* — African Proverb

Remember: Baby steps still move you forward!

## Evaluation Time! (10 minutes)

- Post-Test
- Workshop evaluation...tells us about pace, style of presenting, etc.

## In 3 Months...

- Emailing you a brief survey around your 3-3-3 Action Plan
- Google Hangout?

**Please note: The NEH requires us to do these things...and it helps to make sure these workshops are delivering outcomes that bring tangible results to our peers!**

Gauge to see if there is interest in a Google hangout in 3 months.

# Thank You for Coming!

## PLEASE RETURN:

- Pre & Post Tests
- Workshop Evaluation



Insert NWA and SHN logos

# POWRR Project Team Members

Contact us...we are here to help!

## *Northern Illinois University*

**Lynne M. Thomas** Curator, RBSC  
[lmthomas@niu.edu](mailto:lmthomas@niu.edu) 815.753.0255

**Drew VandeCreek** Director Digital  
[drew@niu.edu](mailto:drew@niu.edu) Scholarship  
815.753.7179

**Jaime Schumacher** Scholarly Communications  
[jschumacher@niu.edu](mailto:jschumacher@niu.edu) Librarian  
815.753.0576

**\*Danielle Spalenka** Digital POWRR Director  
[powrr@niu.edu](mailto:powrr@niu.edu)

**\*Stacey Erdman** Digital Collections Curator  
[serdman@niu.edu](mailto:serdman@niu.edu) 815.753.1004

## *Chicago State University*

**Aaisha Haykal** University Archivist  
[anhaykal@gmail.com](mailto:anhaykal@gmail.com)

**\*Martin Kong** Systems Librarian  
[martinkong2@gmail.com](mailto:martinkong2@gmail.com)

## *Illinois State University*

**Patrice-Andre Prud'homme** Digital Collections Head  
[ppprudh@ilstu.edu](mailto:ppprudh@ilstu.edu) 309.438.5385

## *Illinois Wesleyan University*

**Meg Miner** University Archivist  
[mminer@iwu.edu](mailto:mminer@iwu.edu) 309.556.1538

## *Western Illinois University*

**\*Jeff Hancks** Director, Archives and  
[jl-hancks@wiu.edu](mailto:jl-hancks@wiu.edu) Special Collections  
309.298.2717

\* Presented today