OSSArcFlow
Researching Archival Workflows for Born-Digital Content

Sam Meister
@samalanmeister
POWRR Institute
April 18, 2019
#ossarcflow | ossarcflow_public@educopia.org

EDUCOPIA INSTITUTE

INSTITUTE of Museum and Library SERVICES

A **single turnkey system** does not exist and is unlikely to exist in the future

Why open source software?

How can institutions

combine tools

to support ***workflows***

that meet local institutional needs?

How can institutions implement ***“handoffs”*** between different function-based systems?

What socio-technical factors and

institutional drivers

influence decision making

as institutions choose tools and create workflows?

investigate

model

test

12

partner institutions

Atlanta University Center, Robert W. Woodruff Library
District of Columbia Public Library
Duke University
Emory University
Kansas Historical Society
Massachusetts Institute of Technology
Mount Holyoke College
New York Public Library
New York University
Odum Institute
Rice University
Stanford University

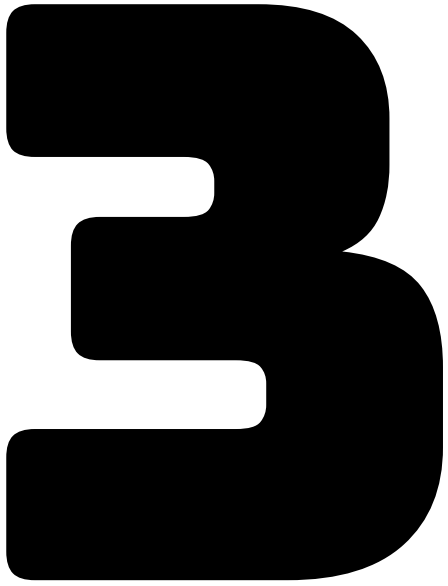
partner needs

Identify and test tools that will help us to implement our digital preservation plan

Collaborate with experts in the field and other institutions grappling with similar issues

Move towards greater consistency in our workflows across collections

Ways to automate and streamline existing digital curation workflows



open source software environments


BitCurator


ArchivesSpace
a community served by LYRASIS


archivematica

10

project team members

Christopher “Cal” Lee (UNC SILS)

Katherine Skinner (Educopia)

Sam Meister (Educopia)

Jessica Meyerson / Alex Chassanoff (Educopia)

Courtney Vukasinovic / Caitlin Perry (Educopia)

Kam Woods (UNC SILS)

Andrew Rabkin (UNC SILS)

Yinglong Zhang (UNC SILS)

Colin Post (UNC SILS)

Kelly Stewart, Sarah Romkey (Artefactual)

Laney McGlohon / Christine DeBella (ArchivesSpace)

document workflows

develop tools

create implementation guide

document workflows

Surveys
Semi-structured interviews
Visual Modeling

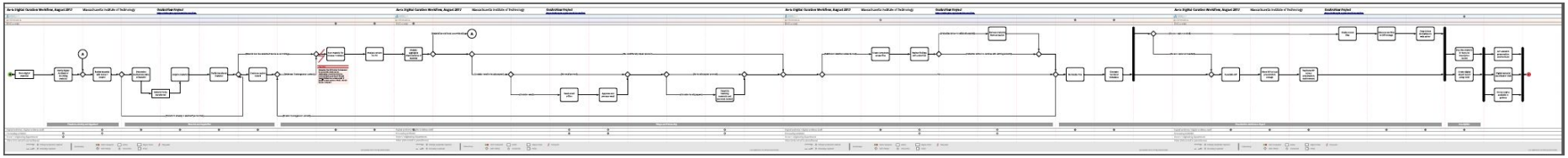
represent workflows

Procedural narratives

Tabular steps

Visual diagrams

step	description	software	hardware
Notify Digital Archivist of incoming material	[Digital Archivist] is called in on any digital acquisition		
Review material with donor / creator	[Digital Archivist] determines and provides donor services so that the donor can review and appraise their content to determine what they want to donate		
Review material with donor / creator	[Digital Archivist] works with the department to determine: Do they have a private server in the department or lab? Do they have content on media (jump drives, disks, etc.)? Do they use the institutional email server? Is there a shared server space that MIT manages that they keep files on? Do they have any files on a commercial cloud service provider? Who has access to that content that can provide [Digital Archivist] with access to that content?		
Determine method and date of transfer	[Digitial Archivist] determines method and date of transfer		
Verify transferred material	[Digital Archivist] verifies material transferred based on log		
Create accession record	[Digitial Archivist] creates an accession stub record, gets accession number		



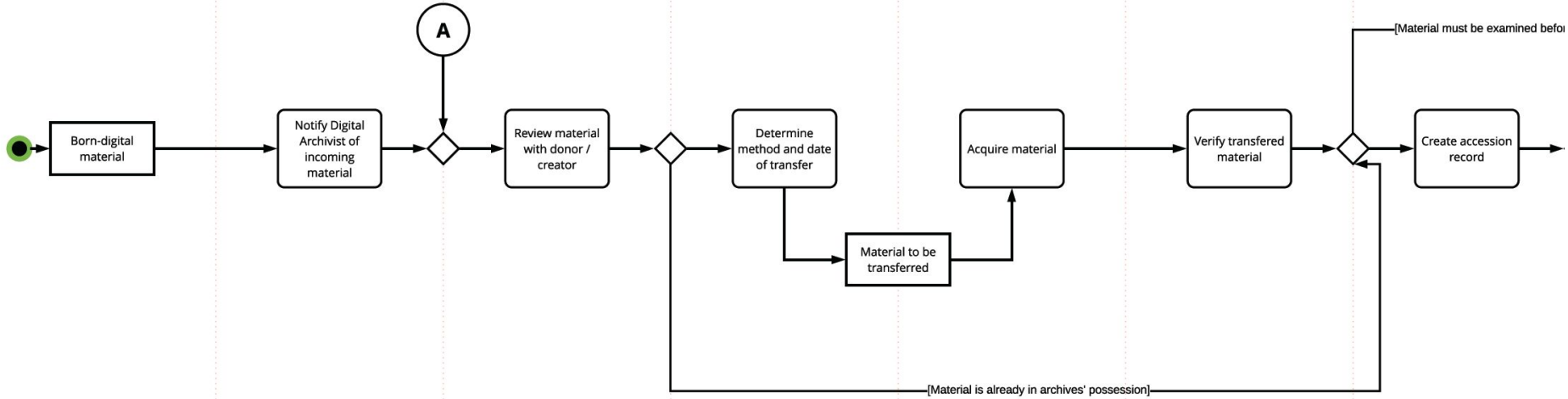
As-Is Digital Curation Workflow, August 2017 Massachusetts Institute of Technology ***OssArcFlow Project***

https://educopia.org/research/ossarcflow

ArchivesSpace

Archivematica

BitCurator



	Pre-Accessioning and Appraisal		Transfer and Acquisition			
Digital archivist / digital archives staff		○	●	●	●	●
Processing archivist	○	○				
Donor / originating department	○					
Other (role noted in parentheses)						

Primary / preferred / required — Secondary / optional — Concurrency — Start / end point — Split / Merge — Action — Connection — Object / Data — Delay — Pain point

pain points

“It would be nice if there was a way for BCE to talk to Aspace. But it’s a quarantine system. Or if there was a way to track information – a lot of spreadsheets. In Aspace or Archivematica.”

pain points

“I want to update descriptive information and extent in ArchivesSpace based on processing (disk imaging, file extraction, etc.) using BitCurator without having to manually do this.”

develop handoffs

Analyze and synthesize gaps and challenges across workflows

Identify pain points and potential metadata handoffs

develop handoffs

Create development tasks for system hand-offs

Generate scripts to address potential system hand-offs

test

rinse

repeat

create implementation guide

Illuminate the ways that socio-technical factors influence standardized workflow development and tools implementation for born-digital archiving in libraries, archives, and museums

create implementation guide

How-To Build/Improve Your Workflow

How to start and/or improve your born-digital workflow processes (including functions and tools and outputs)

Identify concrete steps found in all partner workflows (range of tools and practices, about 8-10 steps, not necessarily in the same order)

Identify pathways to preservation-readiness

Stay in Touch!

Project Webpage:

educopia.org/research/ossarcflow

OSSArcFlow Public List:

https://groups.google.com/a/educopia.org/d/forum/ossarcflow_public

EDUCOPIA INSTITUTE

INSTITUTE of Museum and Library SERVICES

exercise

Yes, even you can create workflow documentation!

exercise

1. Pair up
2. Take turns responding to each interview question

exercise

1. What is the starting point for your workflow?

exercise

2. What are the main steps / activities in your workflow?

exercise

3. Who are the people / roles involved in your workflow activities?

exercise

4. What systems / tools are used during these main activities?

Learn More!

educopia.org

@Educopia

bitcuratorconsortium.org

@bitconsortium

metaarchive.org

@metaarchive

EDUCOPIA INSTITUTE

INSTITUTE of Museum and Library SERVICES