

Enhancing Services to Preserve New Forms of Scholarship

The Charleston Conference, November 8, 2019

Jonathan Greenberg, NYU Libraries & NYU Press

Charles Watkinson, Michigan Publishing

Karen Hanson, Portico

Susan Doerr, University of Minnesota Press



Project Goals

Led by NYU Libraries and funded by the Andrew W. Mellon Foundation, the project is a collaboration of libraries, preservation organizations, and university presses with 2 goals:

Goal 1: Given samples of the works the preservation service providers will identify what aspects can be preserved, and how scalable/resource-intensive the service would be.

Goal 2: Based on results, and on feedback from publishers and experts in preservation, copyright, and scholarly communications, develop guidelines and best practices that publishers and authors can use to improve preservability of the monographs.

New Forms of Scholarship

... push the boundaries of how scholars communicate and publishers publish.

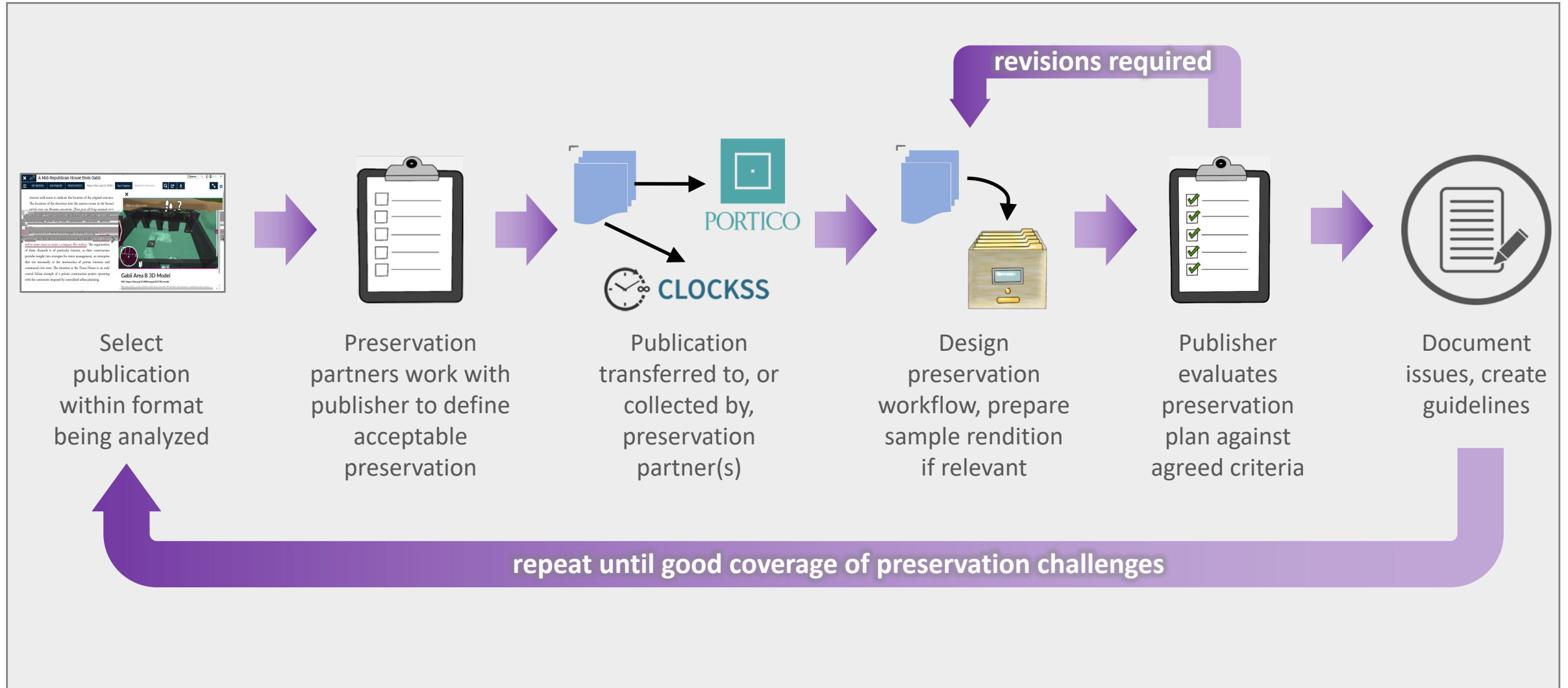
... include **dynamic features** that cannot be represented in print such as:

- embedded audio / video / high-resolution images
- non-linear navigation
- data visualizations
- interactive maps
- supporting software, code, data
- annotations and highlights

... are created **using flexible formats** such as EPUB3 and HTML5. Each depend on **web languages**.

... are a ***preservation challenge!***

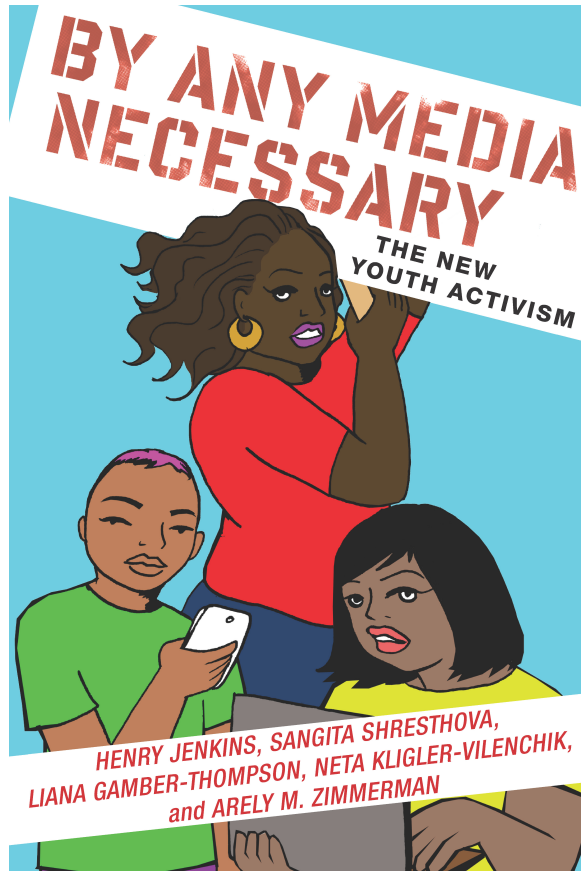
Workflow

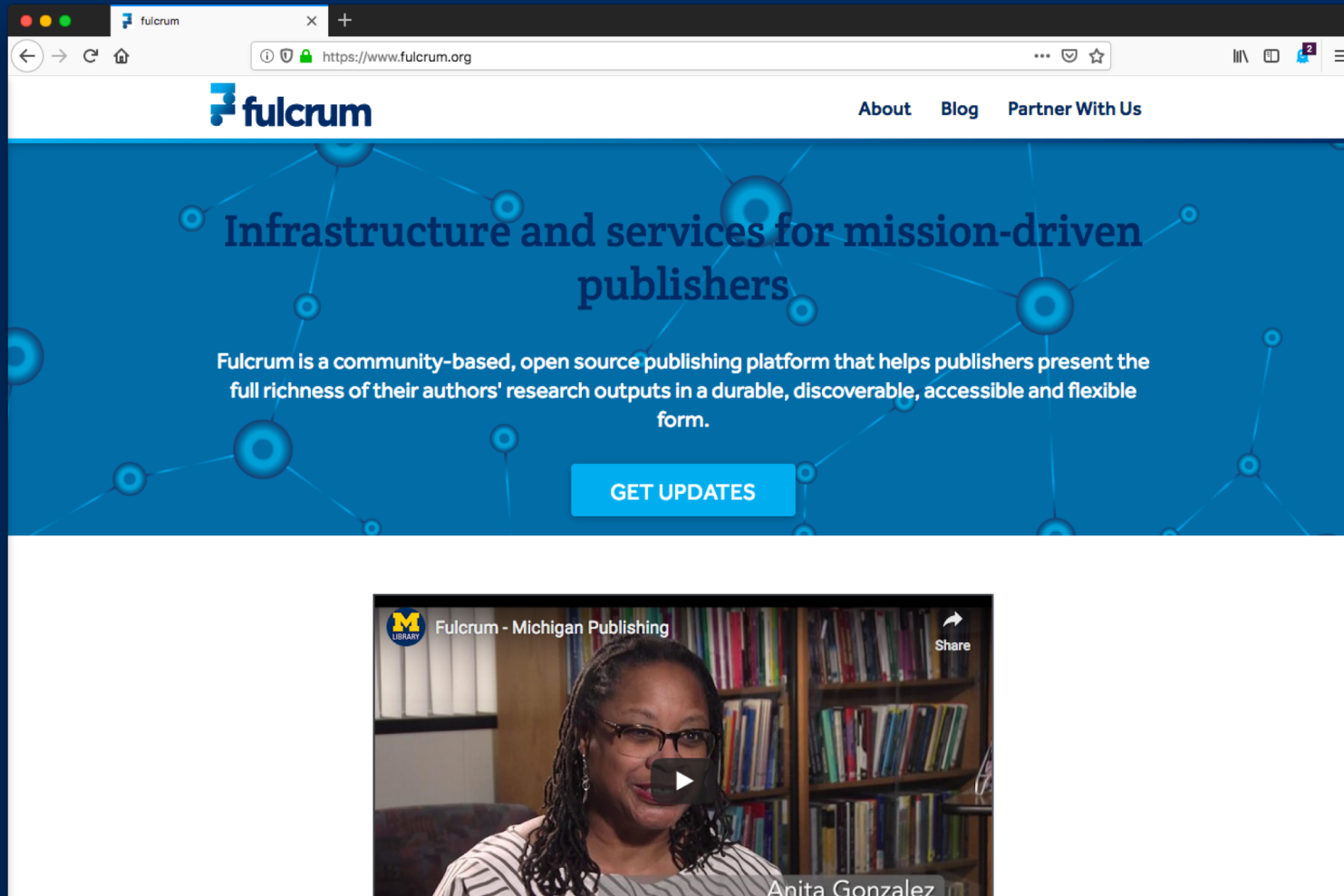


Open Square

- In-house platform for OA books
- Enhanced Ebooks (in collaboration with NYU Digital Library Technology Services)
- Radium browser-based EPUB reader
- New features and design for a rebuilt Radium reader to come

Enhanced Ebooks from NYU Press





Four design principles

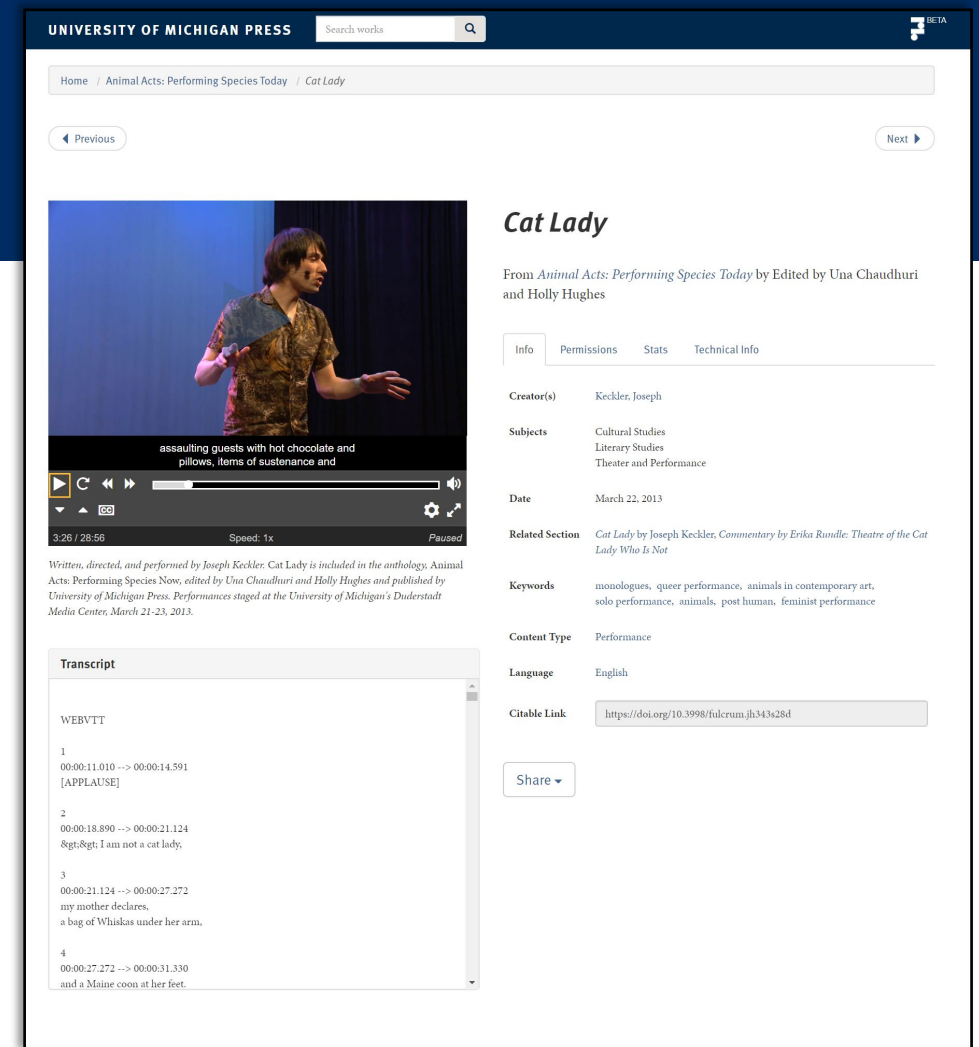
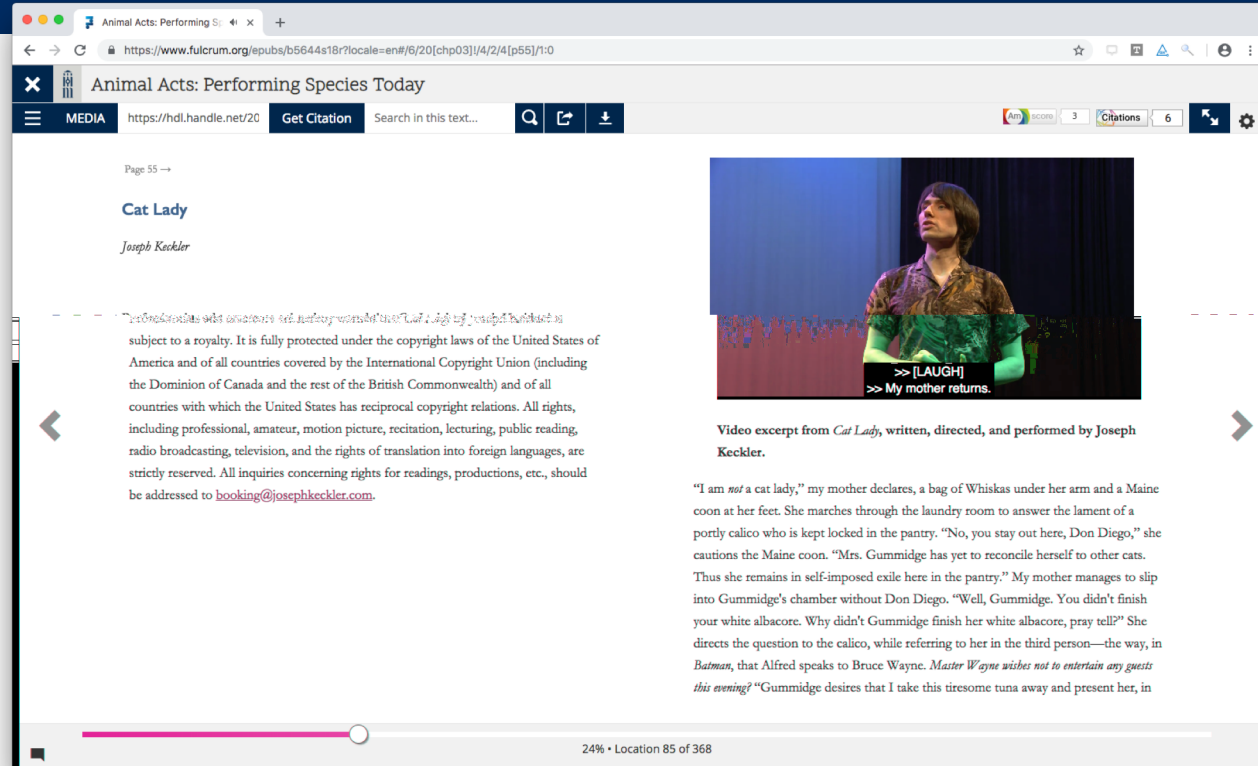
1. Flexible
2. Durable
3. Accessible
4. Discoverable



Fulcrum is an open source publishing platform for long-form digital scholarship. It responds particularly to the increasing need humanists have to integrate their “data” into their “narratives.”



“Enhanced Ebook”



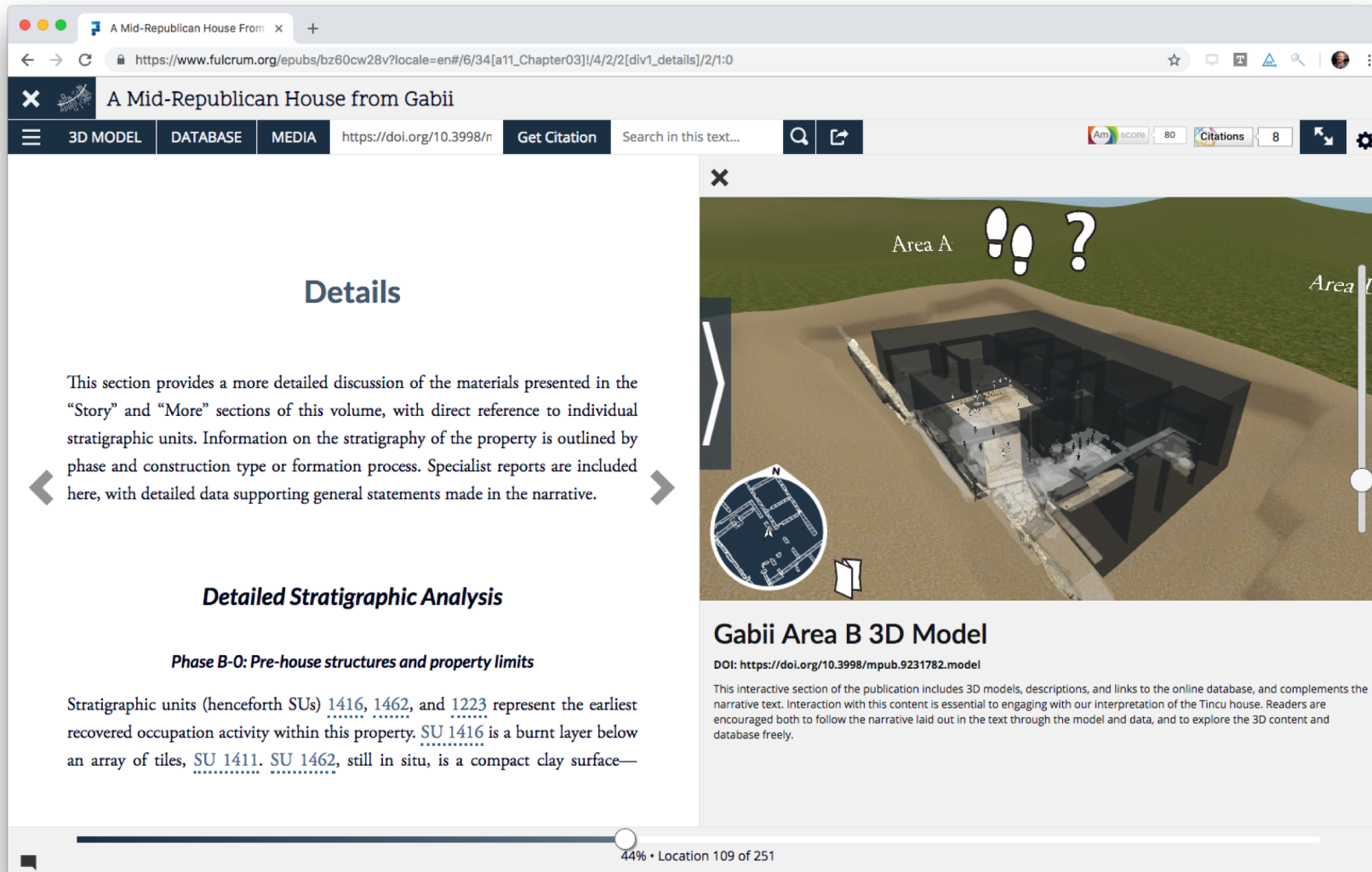
Reading layer:

- EPUB3 opened in EPUB.js reader
- Embedded AblePlayer renders video resource

Repository layer:

- Resource managed in Hyrax repository
- Comprehensive metadata and accessibility
- Preservation pathways for EPUBs/PDFs (to CLOCKSS) and all resources (to APTTrust)

“Interactive scholarly work”



- EPUB3 in EPUB.js
- Unity 3D Model in WebGL
- Media resources managed in Hyrax
- Link out to SQLite database
- Deep two-way linking

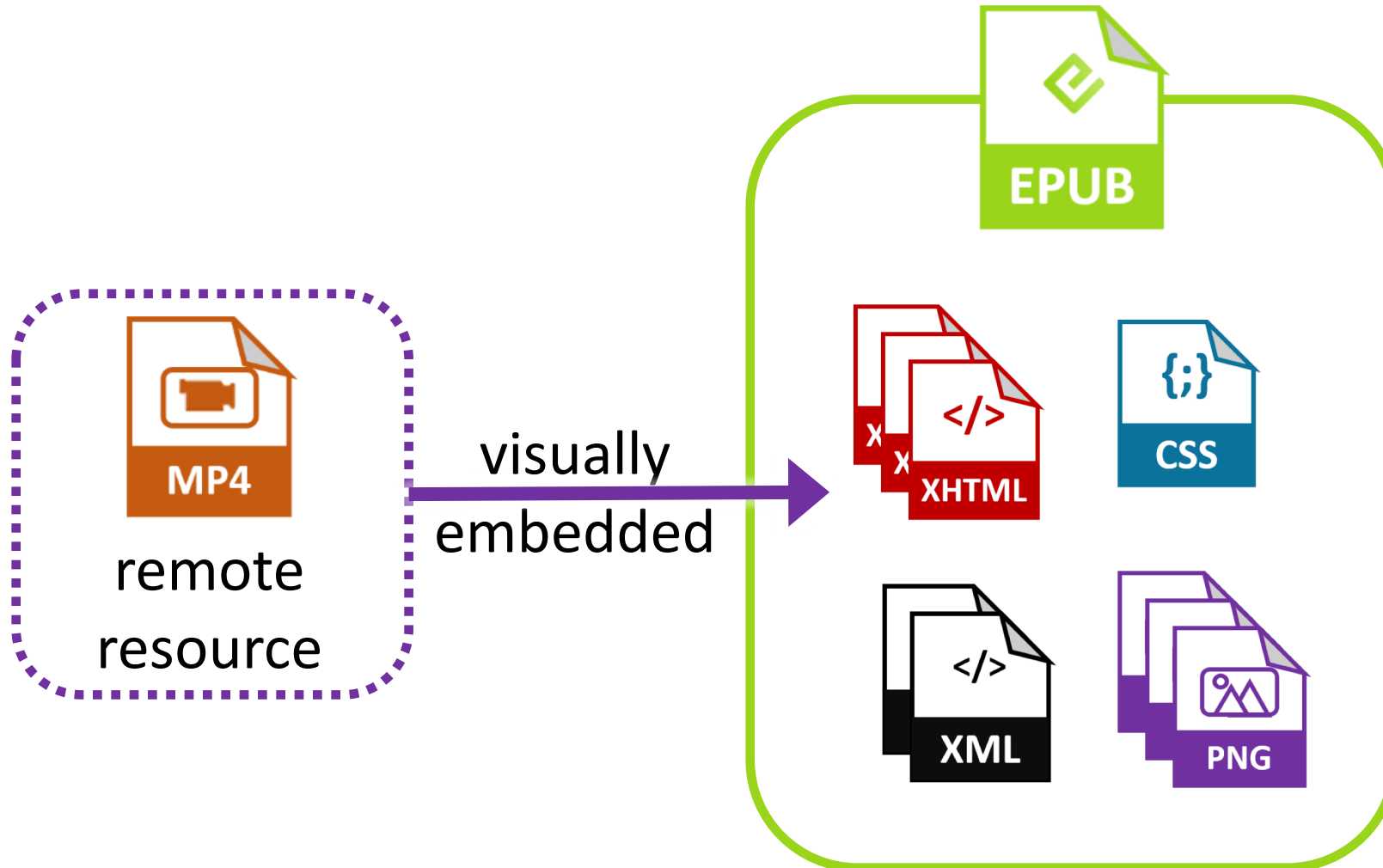
Stewardship commitment from U-M Library. Components can be preserved, but preserving interactivity between them is challenging.

What is Portico?

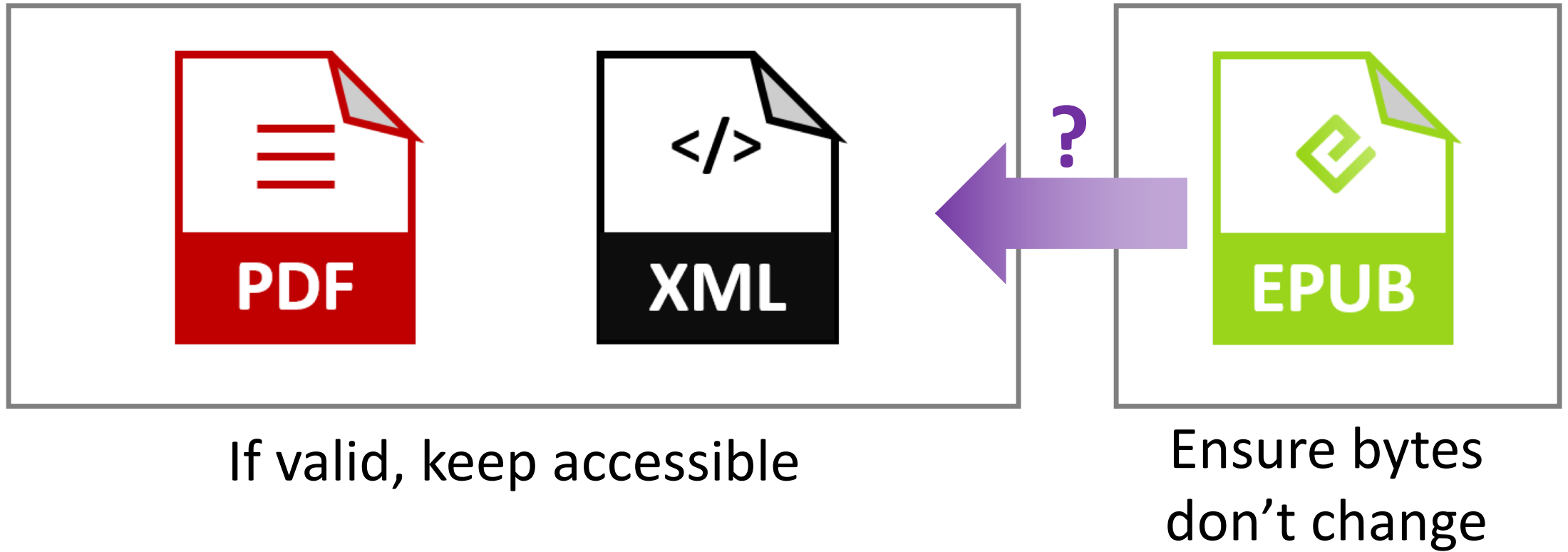
- Community-supported digital preservation archive
- Preserves scholarly publications
- If content no longer hosted online, it becomes available through Portico



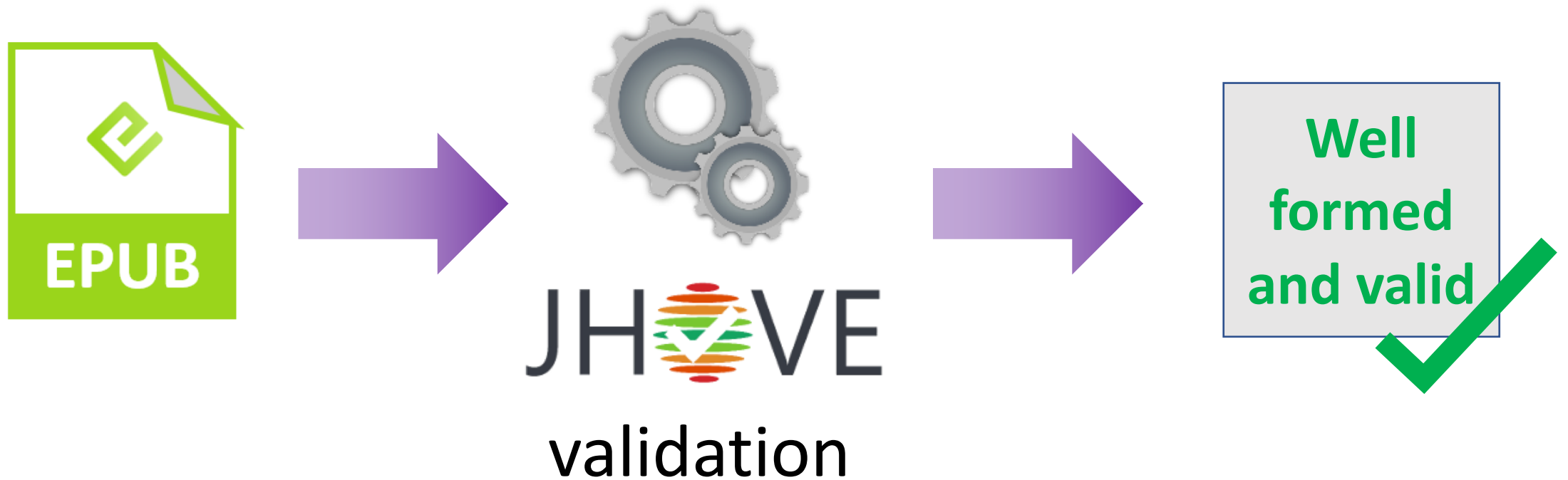
Some EPUB3 tidbits



E-books in Portico



New EPUB JHOVE module



EPUB rendition


PORTICO
Dashboard

🔑 = Perpetual Access
🟢 = Not ingested

DOWNLOAD HOLDINGS

COMPARE HOLDINGS

UNIVERSITY OF MICHIGAN PRESS FULCRUM E-BOOKS

UMP Fulcrum E-Books(01 September 2019)


Animal Acts: Performing Species Today
Una Chaudhuri and Holly Hughes (2014)
ISBN: 9780472029532
DOI: 10.3998/mpub.5633302
Portico Content Set: UMP Fulcrum E-Books
Portico Item ID: ark:/27927/dfdxyz987


Publisher University of Michigan Press
Agreement Name University of Michigan Press E-Book Agreement, Version 1.0 (September 01, 2019)
PCA allowed Yes
Stable URL /stable?au=dfdxyz987

PDFs	EPUB	HTML	Preservation Information	Tree View	PMD	XML
File Name dfdabc123.epub (11126kb) VIEW EPUB DOWNLOAD EPUB						
Format application/epub+zip (Electronic Publication:3)						
Original File Name fulcrum.org.heb-abcd1234\data\heb991234.0001.001.epub						

Examples of EPUB challenges


Challenge 1: Embedded YouTube videos

 **Open Square**
Knowledge without Boundaries



BY ANY MEDIA NECESSARY

its real message—social inequality—through the release of a YouTube video called “The Hunger Games Are Real.”



The video starts with a quote from Haymitch, who advises Katniss and Peeta, contestants in the fictional Hunger Games: “From now on, your job is to be a distraction, so people forget what the real problems are.” The video then shows some of the marketing efforts for *Catching Fire*, highlighting the “hot guys” in the cast, only to interrupt the scene with static and a disruptive black-and-white “rebel message.” Lauren Bird, the spokesperson for the HPA, dressed in an austere black outfit like that of a rebel leader in the film, proclaims:


Enough with the distraction, the Hunger Games are real.






Check it out: In *The Hunger Games* a small portion of the population controls a majority of the wealth. People have full time jobs and still go hungry.

Think it's fiction? Think again.


While the audio details the economic disparities in *Hunger Games*' dystopian country of Panem, the visual shows statistics of real-world economic inequality. For example, the death of Katniss' father in a mining accident is linked to the decline of unionization in the United States; the corrupt justice system in Panem is connected to the fact that African Americans are incarcerated at six times the rate of whites in the U.S. Bird concludes:


Does the *Hunger Games* have hot guys in it? Big fucking deal. It also has something else. Us. People who want justice. President Snows of the world, your reign is coming

BROWSE BOOKS • ABOUT OPEN SQUARE • SEARCH • ACCESSIBILITY 


Approach 1: Keep the EPUB as-is

 **Open Square**
Knowledge without Boundaries

 **NYU PRESS**

BY ANY MEDIA NECESSARY

its real message—social inequality—through the release of a YouTube video called “The Hunger Games Are Real.”

Watch later Share

The video starts with a quote from Haymitch, who advises Katniss and Peeta, contestants in the fictional Hunger Games: “From now on, your job is to be a distraction, so people forget what the real problems are.” The video then shows some of the marketing efforts for *Catching Fire*, highlighting the “hot guys” in the cast, only to interrupt the scene with static and a disruptive black-and-white “rebel message.” Lauren Bird, the spokesperson for the HPA, dressed in an austere black outfit like that of a rebel leader in the film, proclaims:

Enough with the distraction, the Hunger Games are real.






Check it out: In *The Hunger Games* a small portion of the population controls a majority of the wealth. People have full time jobs and still go hungry.

Think it's fiction? Think again.

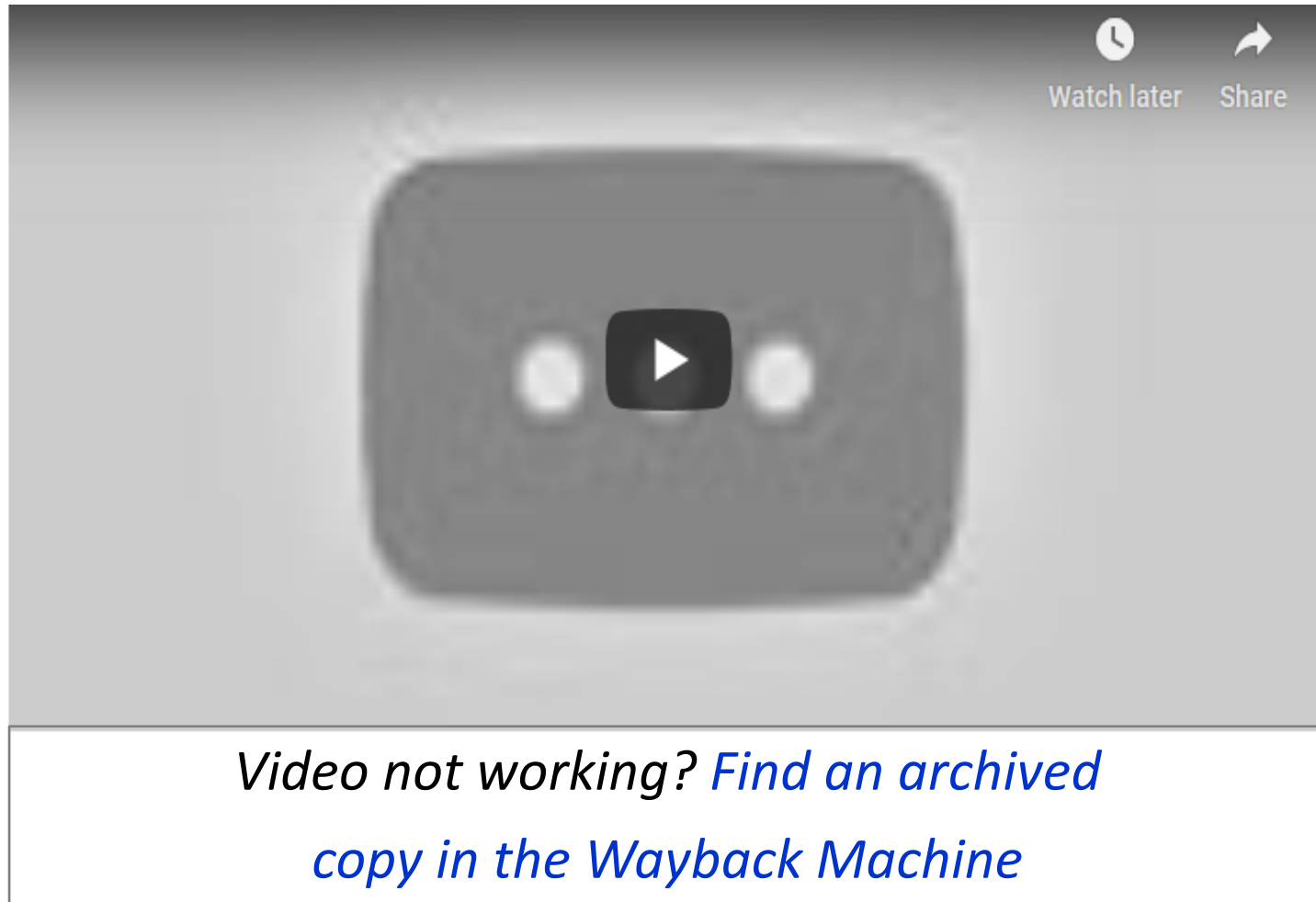
While the audio details the economic disparities in Hunger Games’ dystopian country of Panem, the visual shows statistics of real-world economic inequality. For example, the death of Katniss’ father in a mining accident is linked to the decline of unionization in the United States; the corrupt justice system in Panem is connected to the fact that African Americans are incarcerated at six times the rate of whites in the U.S. Bird concludes:

Does the Hunger Games have hot guys in it? Big fucking deal. It also has something else. Us. People who want justice. President Snows of the world, your reign is coming

[BROWSE BOOKS](#) • [ABOUT OPEN SQUARE](#) • [SEARCH](#) • [ACCESSIBILITY](#)

Approach 2: Automatically save video and link



Copy of video in the Wayback Machine

web.archive.org/web/20190825041326/https://www.youtube.com/watch?v=BmVJaBuoEYA

INTERNET ARCHIVE
Wayback Machine
361 captures
23 Nov 2013 - 25 Aug 2019

https://www.youtube.com/watch?v=BmVJaBuoEYA Go APR AUG 25 SEP 2018 2019 2020

Up next Autoplay ☒

"YOU'VE GOTTA GO"
Ellen Gets Furious & ALMOST Kicks Off Guest!
Plot Twist
358,019 views
12:14

Name: Unknown
Age: 15
Weapon: Unknown
Place: 20th
Death: stabbed in the back by Glimmer (day 1)
6:19

Hunger games ALL tributes info
Kyler
1,570,051 views

5 Harry Potter Theories Too Good Not To Be True
The Why
3,941,899 views
12:47

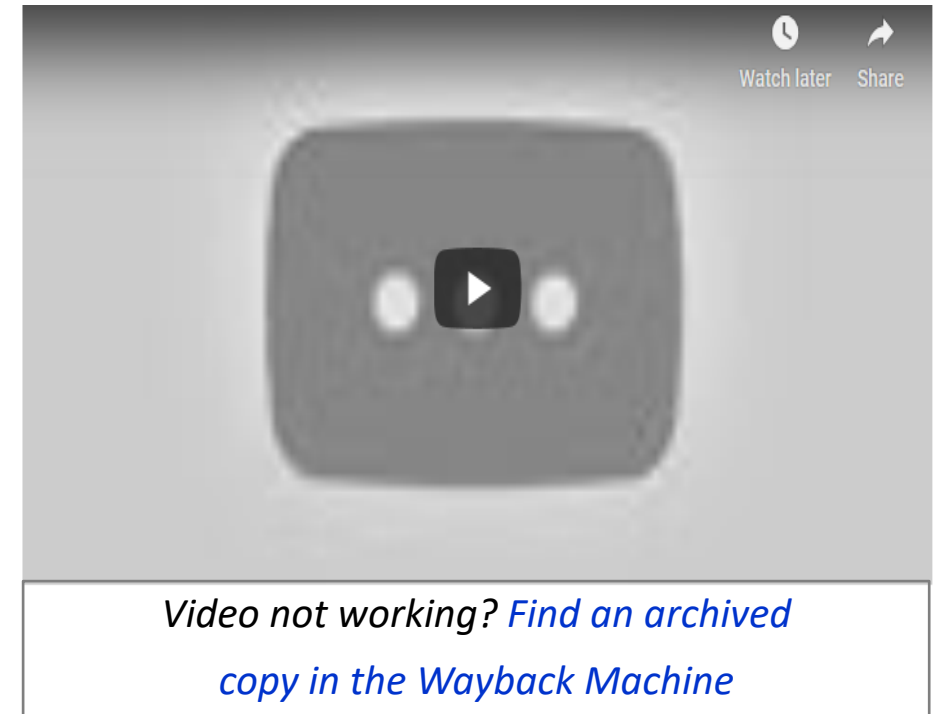
The Life of Helena Ravenclaw (+How to Become a Ghost Explained)
MovieFlame

The Hunger Games Are Real
thehpalliance

Problems with *Portico* taking this approach

- Video may already be gone
- Not always successful, ideally there would be manual QC
- Prefer not to modify content we receive

Ideally this would take place during production of the e-book



Challenge 2: Embedded multimedia viewers

The image shows a web browser window with a video player and a text document. The video player is titled "Cat Lady" and is embedded from "fulcrum.org". It shows a man in a patterned shirt speaking. The text document is titled "Species Today" and contains a citation for the video. A purple arrow points from the video player to the text document.

Species Today

Citation Search in this text... [Search] [Share] [Download]

Am score 3 Citations 6 [Share] [Settings]

Video excerpt from *Cat Lady*, written, directed, and performed by Joseph Keckler.

"I am *not* a cat lady," my mother declares, a bag of Whiskas under her arm and a Maine coon at her feet. She marches through the laundry room to answer the lament of a portly calico who is kept locked in the pantry. "No, you stay out here, Don Diego," she cautions the Maine coon. "Mrs. Gummidge has yet to reconcile herself to other cats. Thus she remains in self-imposed exile here in the pantry." My mother manages to slip into Gummidge's chamber without Don Diego. "Well, Gummidge. You didn't finish your white albacore. Why didn't Gummidge finish her white albacore, pray tell?" She directs the question to the calico, while referring to her in the third person—the way, in *Batman*, that Alfred speaks to Bruce Wayne. *Master Wayne wishes not to entertain any guests this evening?* "Gummidge desires that I take this tiresome tuna away and present her, in

Cat Lady

fulcrum.org/embed?hdl=2027%2Ffulcrum... [Star] [Tools] [K]

1:30 / 28:56 Speed: 1x Playing

24% • Location 85 of 368

Challenge 2: Embedded multimedia widgets



Fig. 3.1. A page from *The Biggest Loser's* website.

The Extra Texts, Bonus Texts, and Ideal Texts of DVDs

If paratexts can brand and recode reality, fictional universes prove an even easier target for branding and recoding. And while fictional films and television shows frequently boast many of the same types of paratexts that makeover shows have, a particularly strong paratext has been the DVD, complete with

them, extra turn that p Thus, for i member to Brookey an package, th writes that meaning a surround" media-indu Bonus mat somewhat are neverth trailers, do audiences v audience m Moreov them and show itself creative pe film's prod film's mean interpretati eggs" can s in fact little the exampl commentar

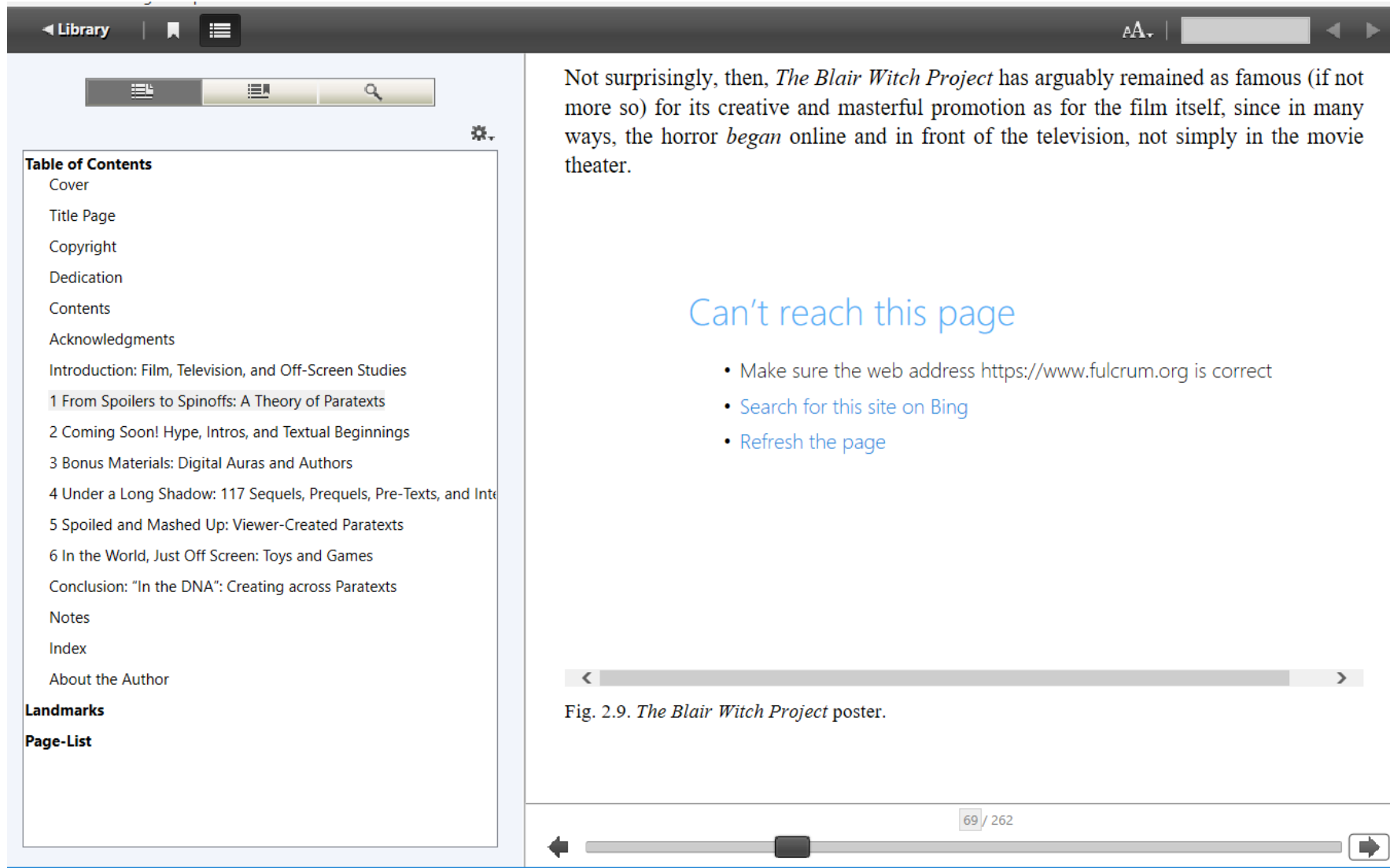


Fulcrum package received by Portico

- The EPUB3 (with remote multimedia viewers)
- Multimedia files, separate from EPUB
- Metadata, including mapping between file names and fulcrum.org web page displayed in EPUB

File Name	Link	Embed Code	Title	Resource Type
l_4.mp4	https://www.fulcrum.com/neverquest-adv	<div style='width:auto;	_Neverquest_ adv	video
l_5.mp4	https://www.fulcrum.com/super-happy-fun	<div style='width:auto;	_Super Happy Fun	video
2_1.jpg	https://www.fulcrum.com/walking-tall-ad	<div style='width:auto;	_Walking Tall_ ad	image
2_2.jpg	https://www.fulcrum.com/baby-mama-ad	<div style='width:auto;	Baby Mama ad	image

Approach 1: Leave EPUB as is



Leaving as-is would require user assembly



Approach 2: Create a large self-contained EPUB

Species Today

Citation Search in this text... [Q] [↗] [↓]

Am score 3 Citations 6 [↗] [⚙]

by warned that *Cat Lady* by Joseph Keckler is
ected under the copyright laws of the United States of
ed by the International Copyright Union (including
est of the British Commonwealth) and of all
ates has reciprocal copyright relations. All rights,
otion picture, recitation, lecturing, public reading,
the rights of translation into foreign languages, are
erning rights for readings, productions, etc., should
[eckler.com](#).

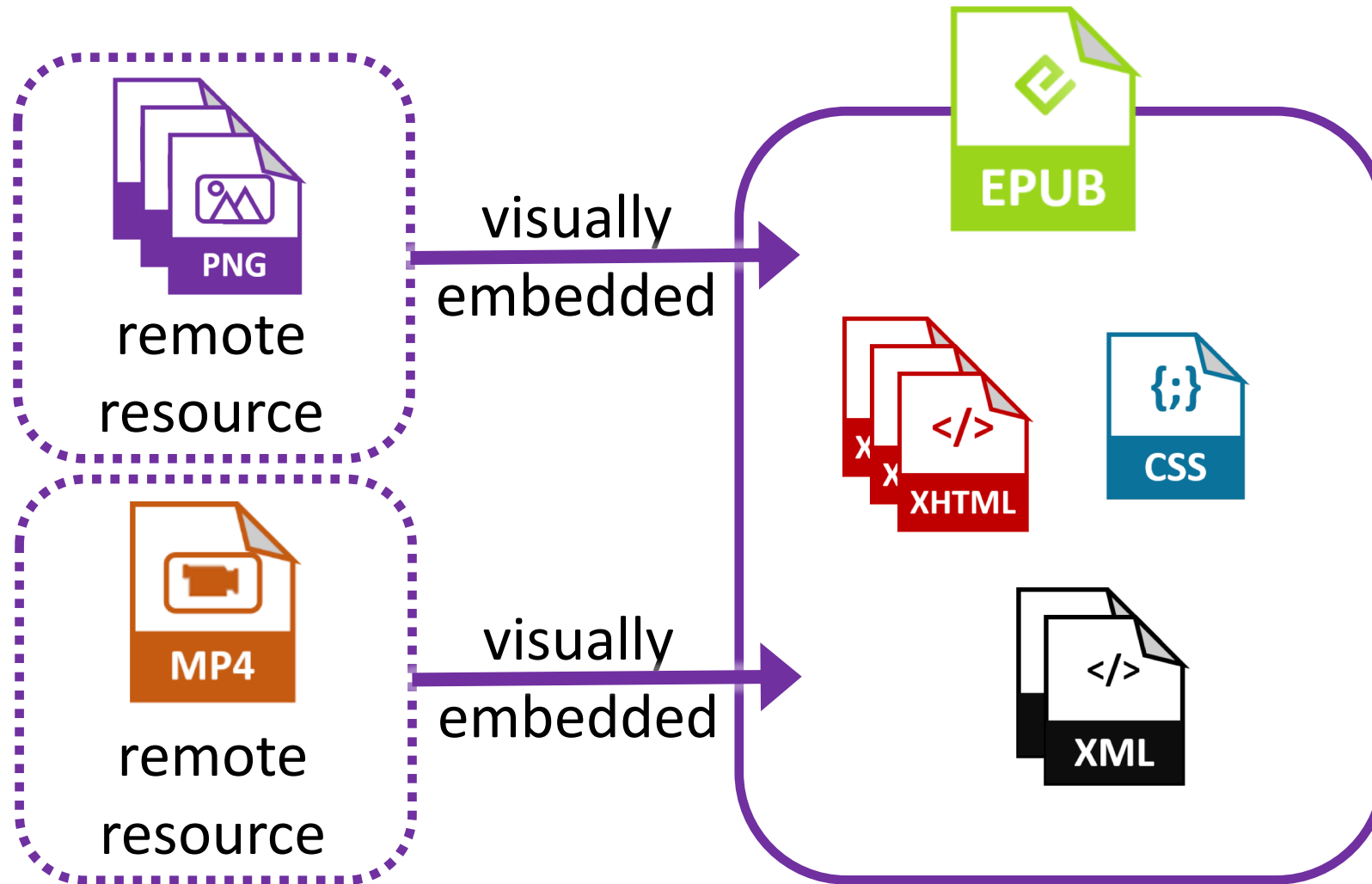
Video excerpt from *Cat Lady*, written, directed, and performed by Joseph Keckler.

“I am *not* a cat lady,” my mother declares, a bag of Whiskas under her arm and a Maine coon at her feet. She marches through the laundry room to answer the lament of a portly calico who is kept locked in the pantry. “No, you stay out here, Don Diego,” she cautions the Maine coon. “Mrs. Gummidge has yet to reconcile herself to other cats. Thus she remains in self-imposed exile here in the pantry.” My mother manages to slip into Gummidge’s chamber without Don Diego. “Well, Gummidge. You didn’t finish your white albacore. Why didn’t Gummidge finish her white albacore, pray tell?” She directs the question to the calico, while referring to her in the third person—the way, in *Batman*, that Alfred speaks to Bruce Wayne. *Master Wayne wishes not to entertain any guests this evening?* “Gummidge desires that I take this tiresome tuna away and present her, in

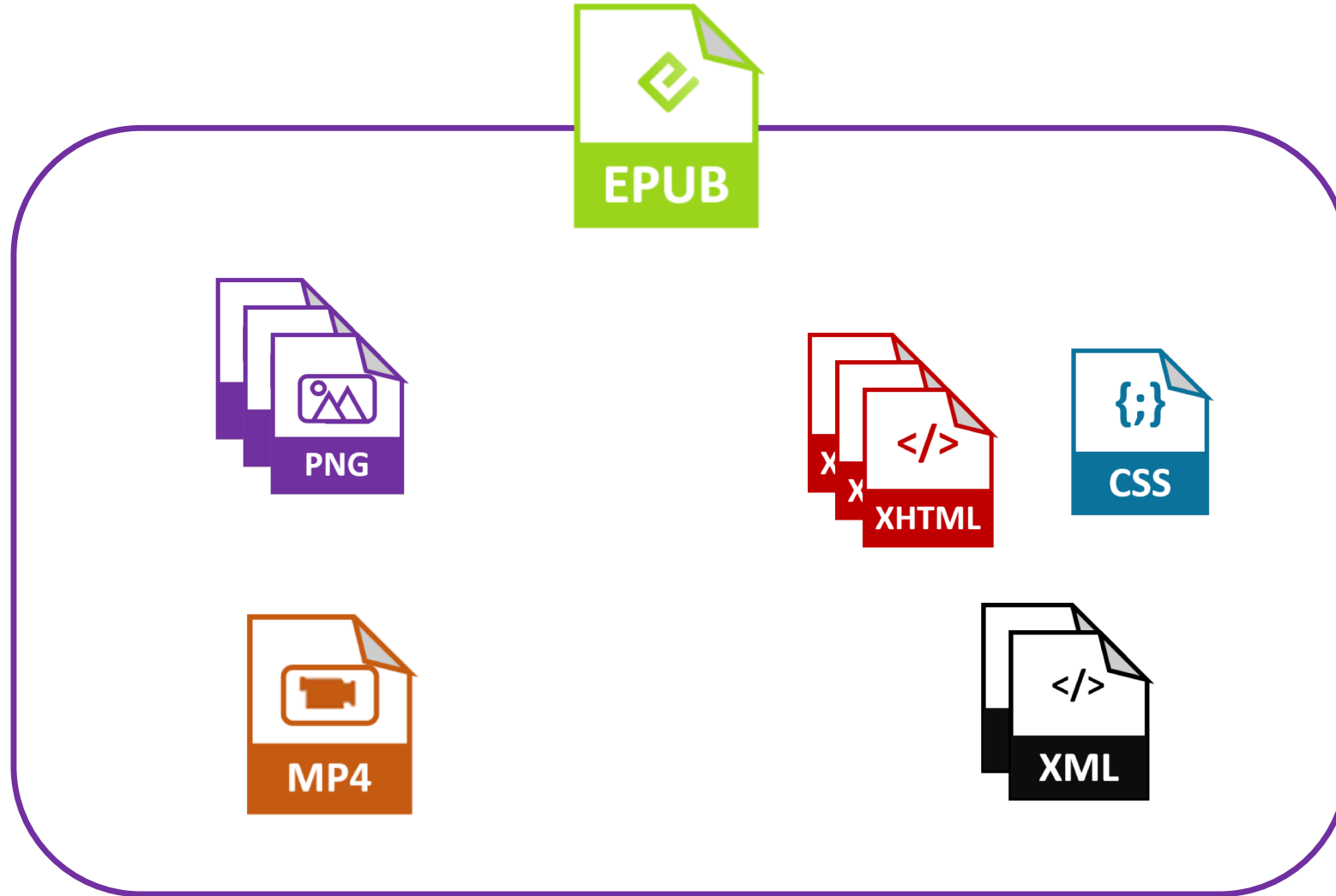
24% • Location 85 of 368

Move remote
resources inside
EPUB

Move the remote resources inside



Move the remote resources inside

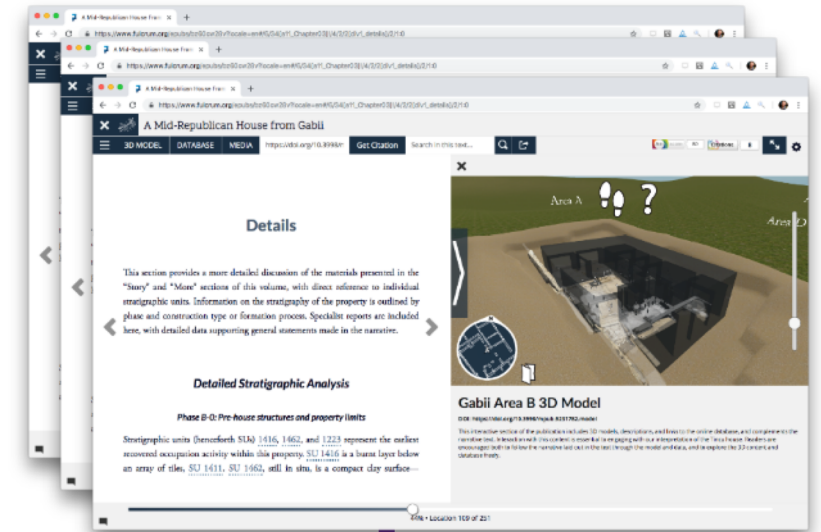


Approach 3: It's behaving like a website, preserve it as a website?

- Would preserve Fulcrum look and feel, including accessibility features

BUT

- It's not an EPUB any more!
- Playback may be more challenging



Harvest website



(web archive file)

All 3 approaches combined?



Original +
supplements



Self-contained
"EPUB/A"??



Web archive of
book on
fulcrum.org

Introducing **Manifold**



Manifold is an intuitive, collaborative, open source web platform for scholarly publishing. With iterative texts, powerful annotation tools, rich media support, and robust community dialogue, *Manifold* seeks to transform scholarly publications into living digital works.

preservation Format:

- EPUB3

ePub is a standard and there is interest in building preservation specifications into the future versions

ePub can contain most of/all of the elements of a Manifold project

Groups that are unable to produce ePub would be able to create them with Manifold (and potentially sell them)

This is potentially a mechanism to serve a complete Manifold edition for an offline reading experience

REUSE Your Metadata!

- Metadata created for Accessibility is useful for Preservation:
 - Long descriptions
 - Keywords
 - alt-text
 - abstracts

But wAit! There's MOre!

What about projects that do not export neatly into ePub?

Cut/Copy/Paste: Fragments of History
by Whitney Treitten

<https://manifold.umn.edu/projects/cut-copy-paste>

Website preservation approaches

Standard website harvester

- Good for simpler websites.
- May struggle with sites that interact frequently with the server based on user interactions



We're at a bit of a loose end.
Frightfully sorry.

404 Error: Page not found

Rhizome's Webrecorder

Archived page from the "manifold exp" Collection on Webrecorder

File View Help

WR C:\Users\khanson\Documents\Workspaces\JhoveTestResults\format-corpus\manifold-exp-20190415200532.warc Open... Help

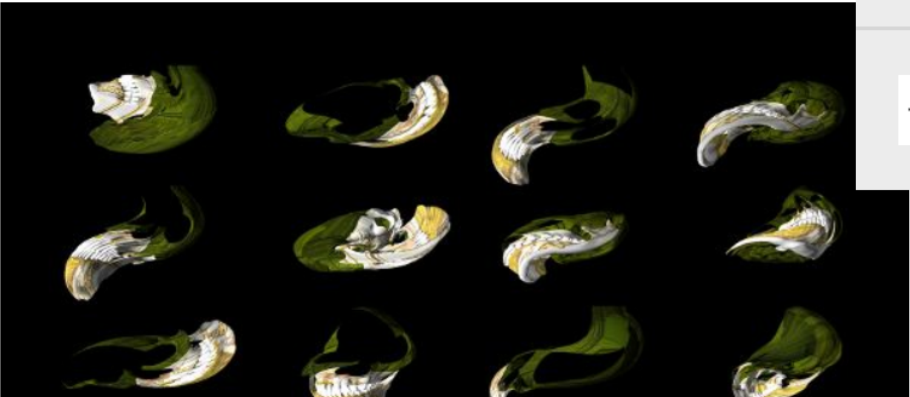
https://manifold.umn.edu/read/metagaming/section/1b1ec89a-49b3-49f5-8678-848a29df8ed1 4/15/2019, 4:04:39 PM

Lists Browse All MENU CONTENTS Metagame 2

Lists (0)

anamorphic distortion. From a randomly generated manifold, multiple levels of digital distortion, the puzzle may have a solution, but it is not for us. It is picture planes all the way down. To solve *Memento Mortem Mortis*, go to <http://manifold.umn.edu/memento-mortem-mortis>

Two small gray cube icons.



Metadata

PAGE PROPERTIES

PAGE TITLE

"Acknowledgments" in "Metagaming" on Manifold

@uminnpress

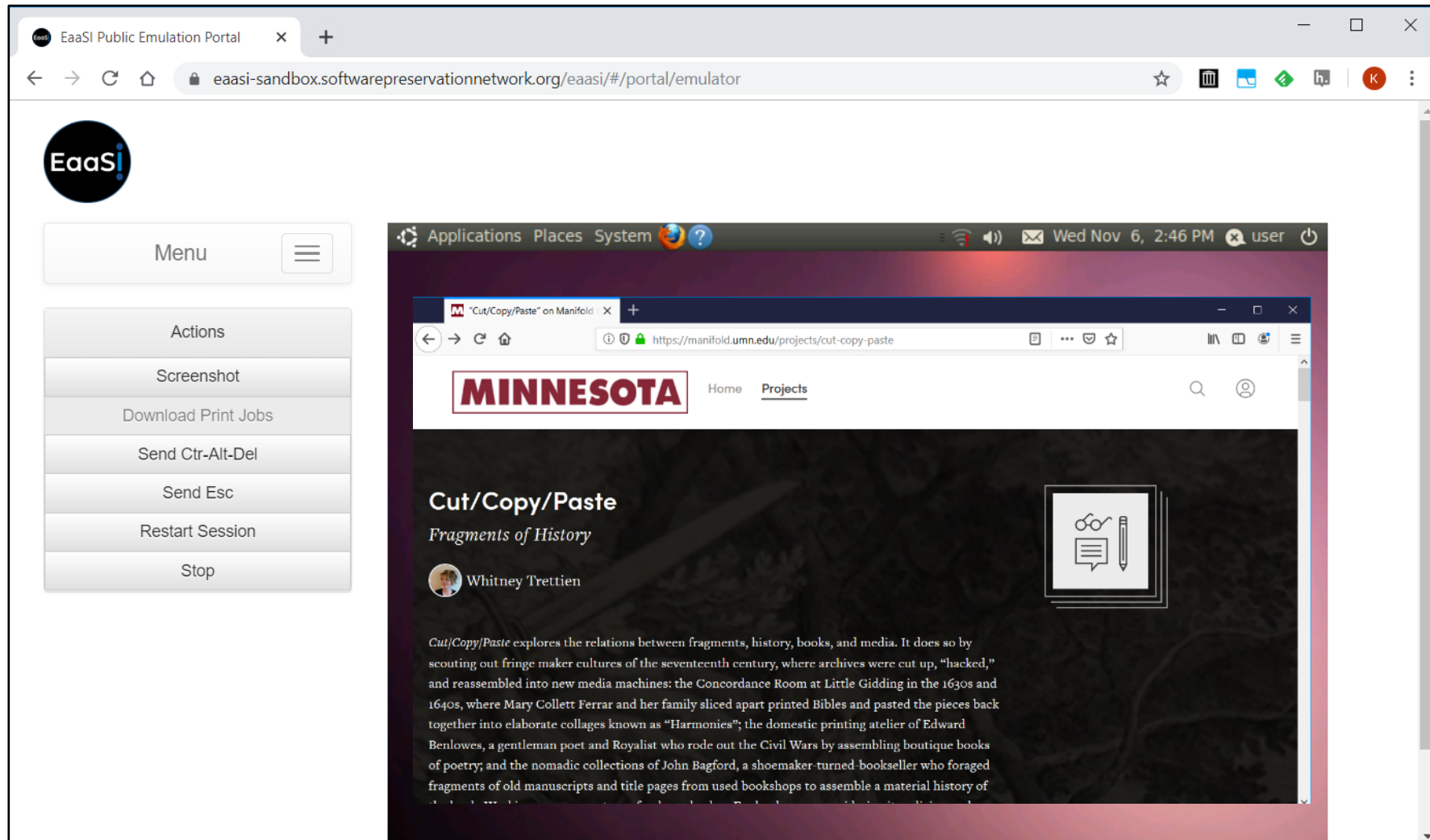
Font settings: Serif, Sans-serif, and a color selection bar.

Something in-between?

- Webrecorder *Autopilot* (Rhizome)
- Memento Tracer (Los Alamos National Lab)

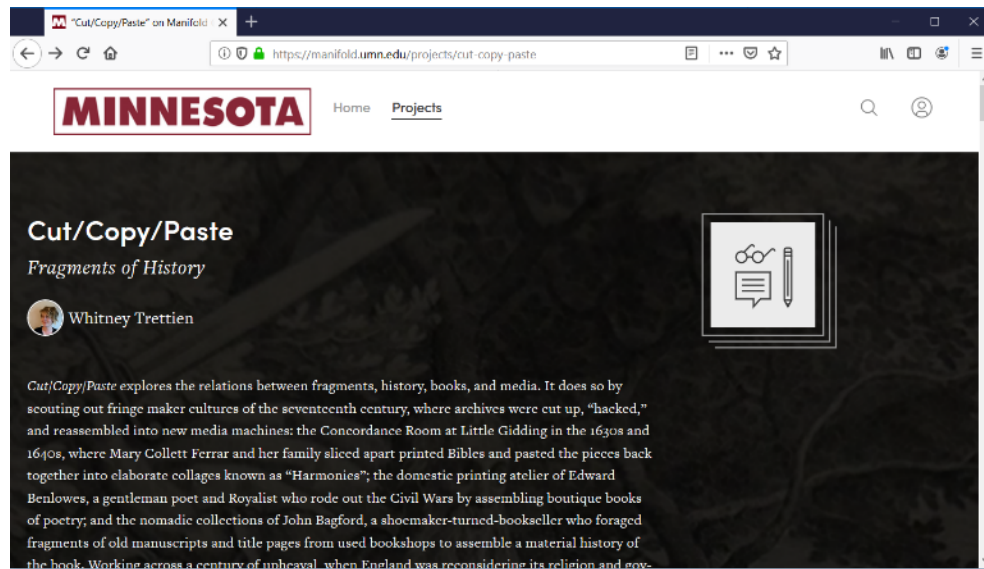
The screenshot displays a web browser window with a URL bar showing a manifold.umn.edu link. The page features a sidebar menu on the left with the following items: "Introducing Edward Benlowes" (highlighted in green), "Mapping Early Modern Libraries", "Collaboration and the Cambridge Press", "Toward a Poetics of the Codex", and "Assembling Theophila". The main content area shows a text excerpt with a green cube icon and the text "More than a secretar...". An "ANNOTATIONS" panel is open on the right, displaying a quote: "Gently stretching our own sense of the verb 'publish' to better match the more flexible early modern use, we might imagine this donation as Brent Hall's first major publication: it was a means of making...". Below the quote are links for "READ FULL PASSAGE" and "LOGIN TO ANNOTATE". A user profile for "Megan Heffernan" (5 months ago) is shown with a comment: "i love this account of publication with an eye to the broader network of reception, one that expands beyond the book trade to fold in library communities. tara lyons' new work on the bodleian's copy of jonson's missing epigrams book is following similar paths." and a "LOGIN TO REPLY" link. At the bottom, a "Doodle of Johann" is partially visible.

Emulation as a Service Infrastructure (EaaSI)



<https://www.softwarepreservationnetwork.org/eaasi/>

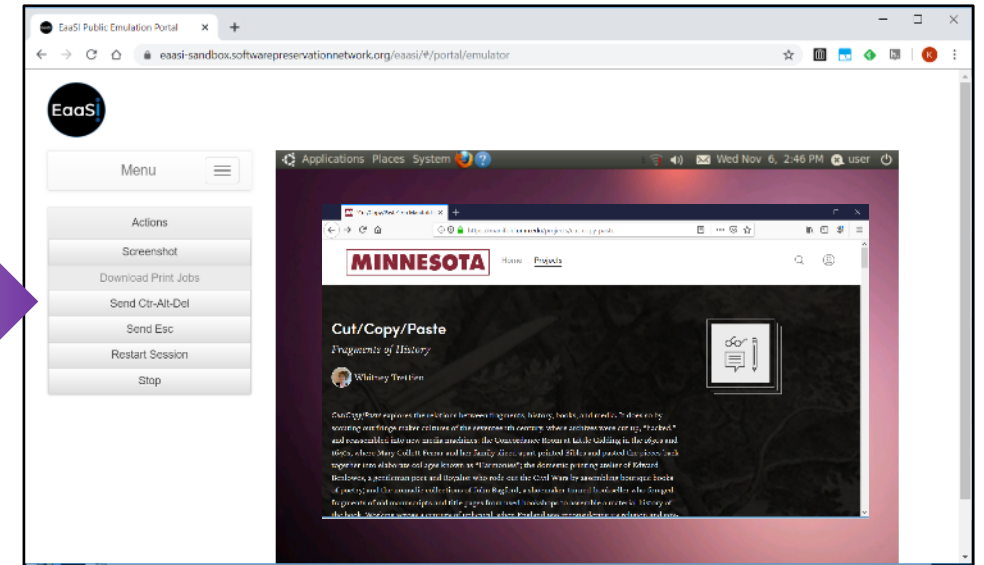
Emulation as a Service Infrastructure (EaaSI)



Live site
@ manifold.umn.edu

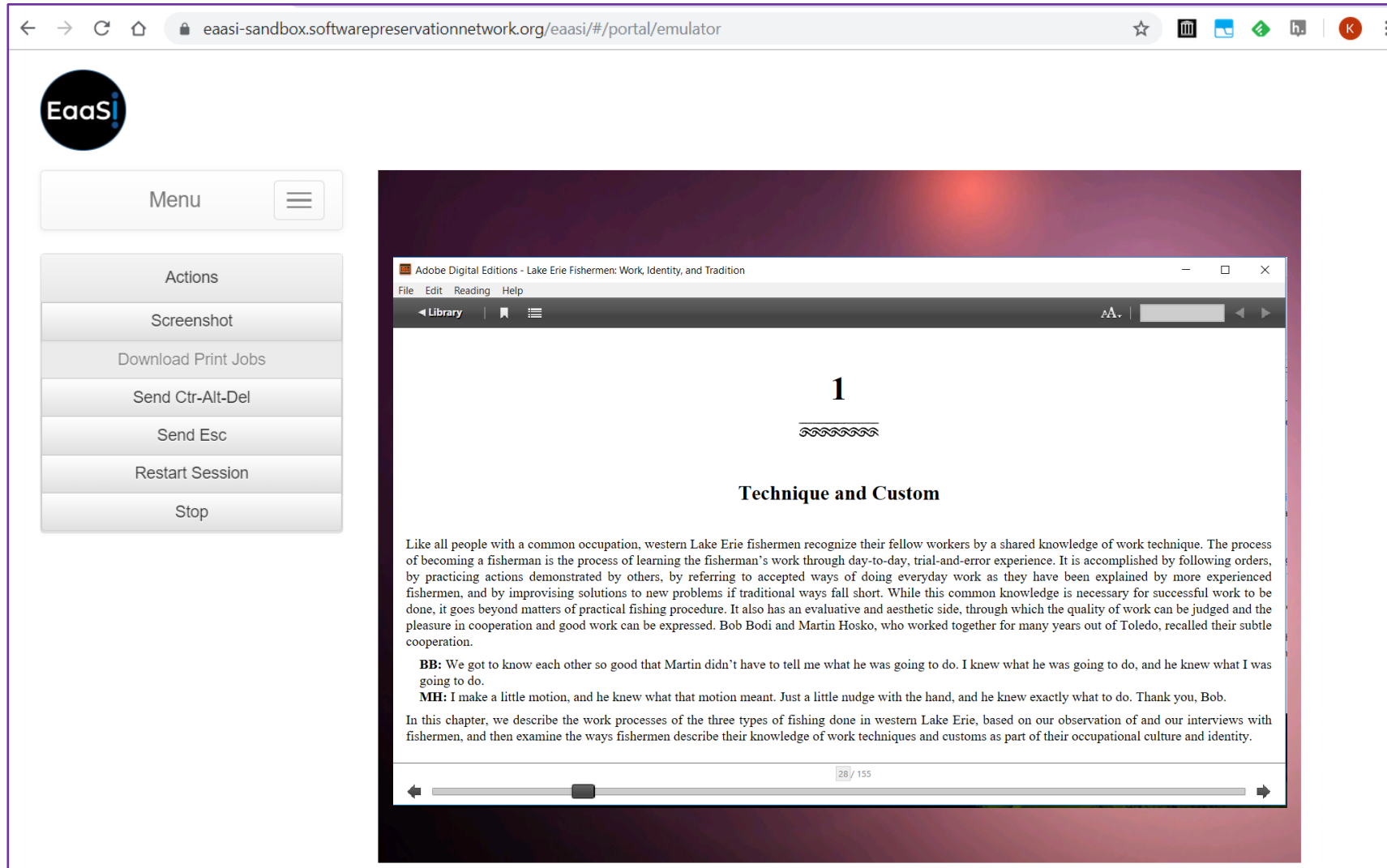


e-book
data



Emulated Manifold
using EaaSI

Eaasi for EPUBs



The screenshot displays the Eaasi web interface within a browser window. The address bar shows the URL: `eaasi-sandbox.softwarepreservationnetwork.org/eaasi/#/portal/emulator`. The interface includes a sidebar on the left with the Eaasi logo and a menu containing the following options: Menu, Actions, Screenshot, Download Print Jobs, Send Ctr-Alt-Del, Send Esc, Restart Session, and Stop. The main content area features a large, dark-themed EPUB viewer window. The viewer's title bar reads "Adobe Digital Editions - Lake Erie Fishermen: Work, Identity, and Tradition". The document content is displayed on a white background and includes the following text:

1

Technique and Custom

Like all people with a common occupation, western Lake Erie fishermen recognize their fellow workers by a shared knowledge of work technique. The process of becoming a fisherman is the process of learning the fisherman's work through day-to-day, trial-and-error experience. It is accomplished by following orders, by practicing actions demonstrated by others, by referring to accepted ways of doing everyday work as they have been explained by more experienced fishermen, and by improvising solutions to new problems if traditional ways fall short. While this common knowledge is necessary for successful work to be done, it goes beyond matters of practical fishing procedure. It also has an evaluative and aesthetic side, through which the quality of work can be judged and the pleasure in cooperation and good work can be expressed. Bob Bodi and Martin Hosko, who worked together for many years out of Toledo, recalled their subtle cooperation.

BB: We got to know each other so good that Martin didn't have to tell me what he was going to do. I knew what he was going to do, and he knew what I was going to do.

MH: I make a little motion, and he knew what that motion meant. Just a little nudge with the hand, and he knew exactly what to do. Thank you, Bob.

In this chapter, we describe the work processes of the three types of fishing done in western Lake Erie, based on our observation of and our interviews with fishermen, and then examine the ways fishermen describe their knowledge of work techniques and customs as part of their occupational culture and identity.

The viewer interface includes a top menu bar with "File", "Edit", "Reading", and "Help", a "Library" sidebar, and a bottom progress bar indicating the current page is 28 of 155.

Thank You & Questions