

4-9-2024

## Will our future selves thank us? An examination of born-digital curation practices at the University of Kentucky Libraries

Ruth E. Bryan

*University of Kentucky Libraries, ruth.bryan@uky.edu*

Sarah Dorpinghaus

*University of Kentucky*

Megan M. Mummey

*University of Kentucky*

Emily B. Collier

*University of Kentucky Libraries*

Andrew McDonnell

*University of Kentucky Libraries, mcdonnell@uky.edu*

Follow this and additional works at: [https://uknowledge.uky.edu/libraries\\_present](https://uknowledge.uky.edu/libraries_present)



Part of the [Archival Science Commons](#)

**Right click to open a feedback form in a new tab to let us know how this document benefits you.**

---

### Repository Citation

Bryan, Ruth E.; Dorpinghaus, Sarah; Mummey, Megan M.; Collier, Emily B.; and McDonnell, Andrew, "Will our future selves thank us? An examination of born-digital curation practices at the University of Kentucky Libraries" (2024). *Library Presentations*. 263.

[https://uknowledge.uky.edu/libraries\\_present/263](https://uknowledge.uky.edu/libraries_present/263)

This Presentation is brought to you for free and open access by the University of Kentucky Libraries at UKnowledge. It has been accepted for inclusion in Library Presentations by an authorized administrator of UKnowledge. For more information, please contact [UKnowledge@lsv.uky.edu](mailto:UKnowledge@lsv.uky.edu).

# Will our future selves thank us?

Examining born-digital curation practices at UKL

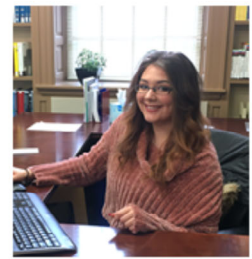
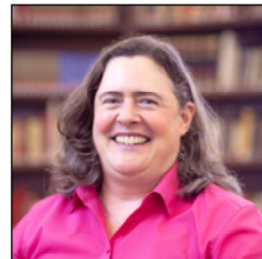


OCLC Works in Progress Webinar series  
April 2024

Hello – I’m Megan Mummey, the Director of Manuscript Collections at the University of Kentucky Libraries. Four of my colleagues and myself will be presenting the panel Will our future selves thank us? An examination of born-digital curation practices at UKL. And quick context here this is an updated version of a presentation that we gave at the Best Practices Exchange Unconference this year. Many of our projects have moved further along and evolved.

## Speakers

- Sarah Dorpinghaus, Director of Digital Strategies and Technology
- Megan Mummey, Director of Manuscript Collections
- Andrew McDonnell, Digital Archivist
- Ruth E. Bryan, CA, University Archivist
- Emily B. Collier, Assistant University Archivist



You will be hearing from us in the following order – First you will hear from Sarah Dorpinghaus about shifting digital preservation infrastructure, then myself on implementing born digital appraisal alongside Andrew McDonnell, then Ruth Bryan on the acquisition of university publications, and then Emily Collier on web preservation. It may seem like we are all talking about disparate subjects, but each presentation will build on each other to form an in-depth case study of how we have been attempting to wrangle the beast that is working with born digital materials.

So if you know me – you know I have a tendency to say flip things (because I'm a youngest child so I'm always

trying to get a laugh). I often say things like “that’s future Megan’s problem”. But I’ve been an archivist for enough time now that I when run across problems, I get angry, and say “who did this!?”...and it’s always “past Megan”. So this panel came together upon the realization that we are all trying to not do this. We are struggling with various pain points, like time, expertise, understaffing, and trying our best to plan for the future in the constantly changing landscape around digital stewardship.



# UKL Digital Preservation Infrastructure

A changing landscape, 2021-2024

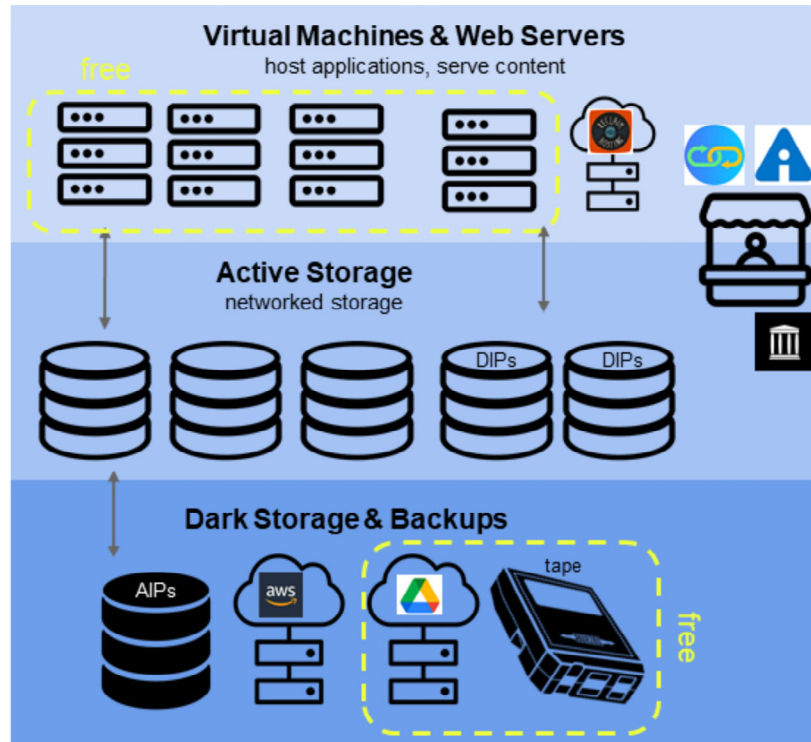


## University of Kentucky Libraries

- R1, land grant institution
- Home-grown digital preservation repository and digital libraries
- UK Libraries has engaged in born-digital archival work since 2015 and web archiving since 2018
- 1 FTE (100%) working with born-digital archives, 8 FTE partial (5-15%), and on average 0-3 student employees

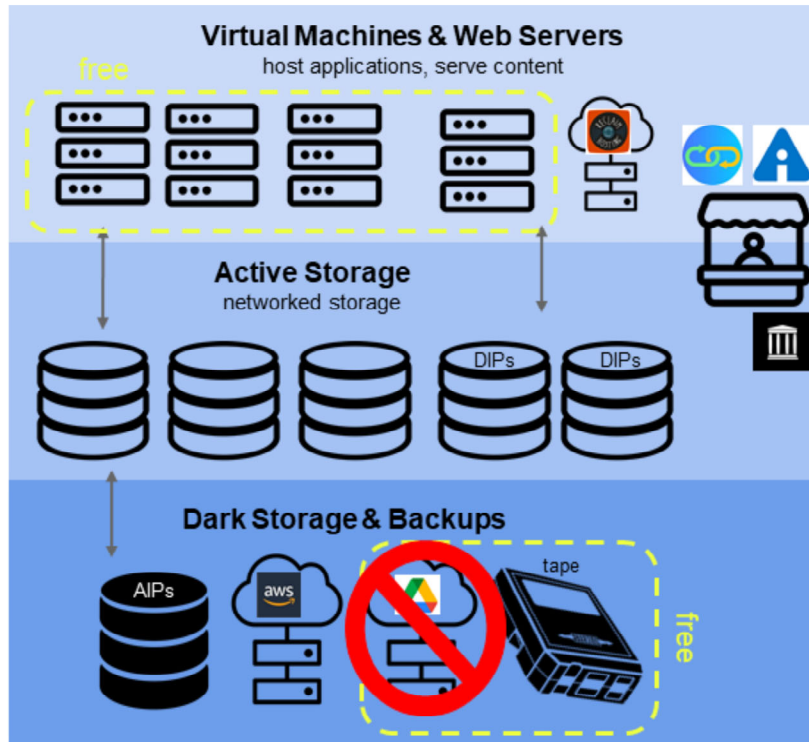


Margaret I. King Library

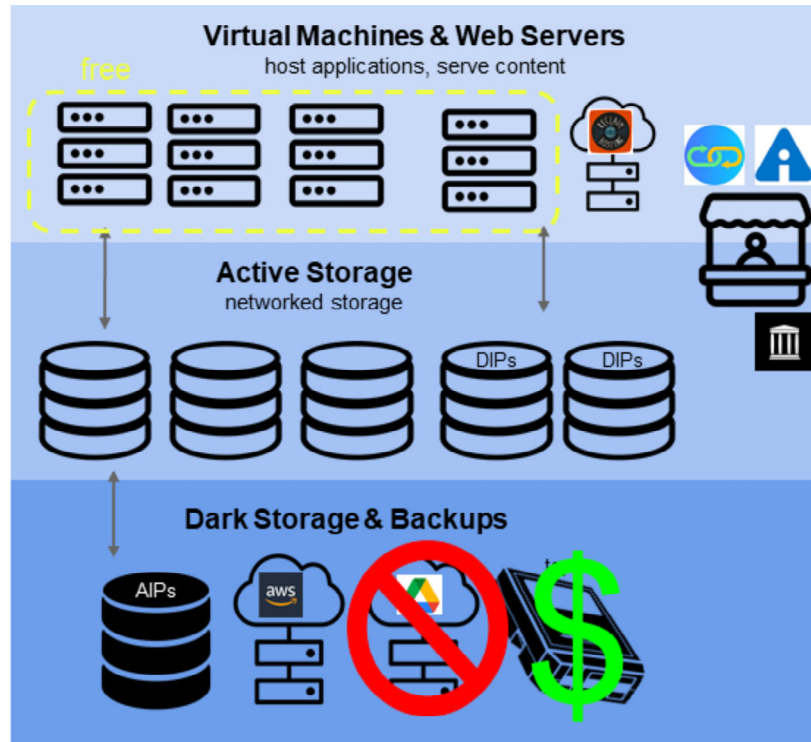


I'll now give a quick overview of our special collections digital infrastructure and some critical changes we've encountered in the past few years and I'll then explain how this has impacted our digital curation decisions.

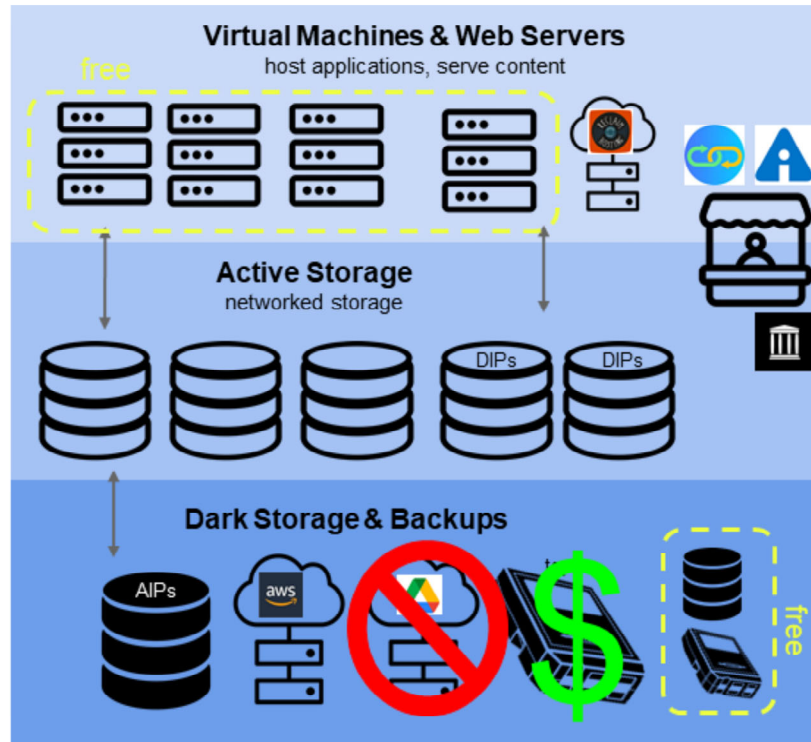
- We have three layers-- dark storage and backups, active networked storage, and VMs and web servers that provide public access to our content, plus some vendor tools like Archive-It and Webrecorder.



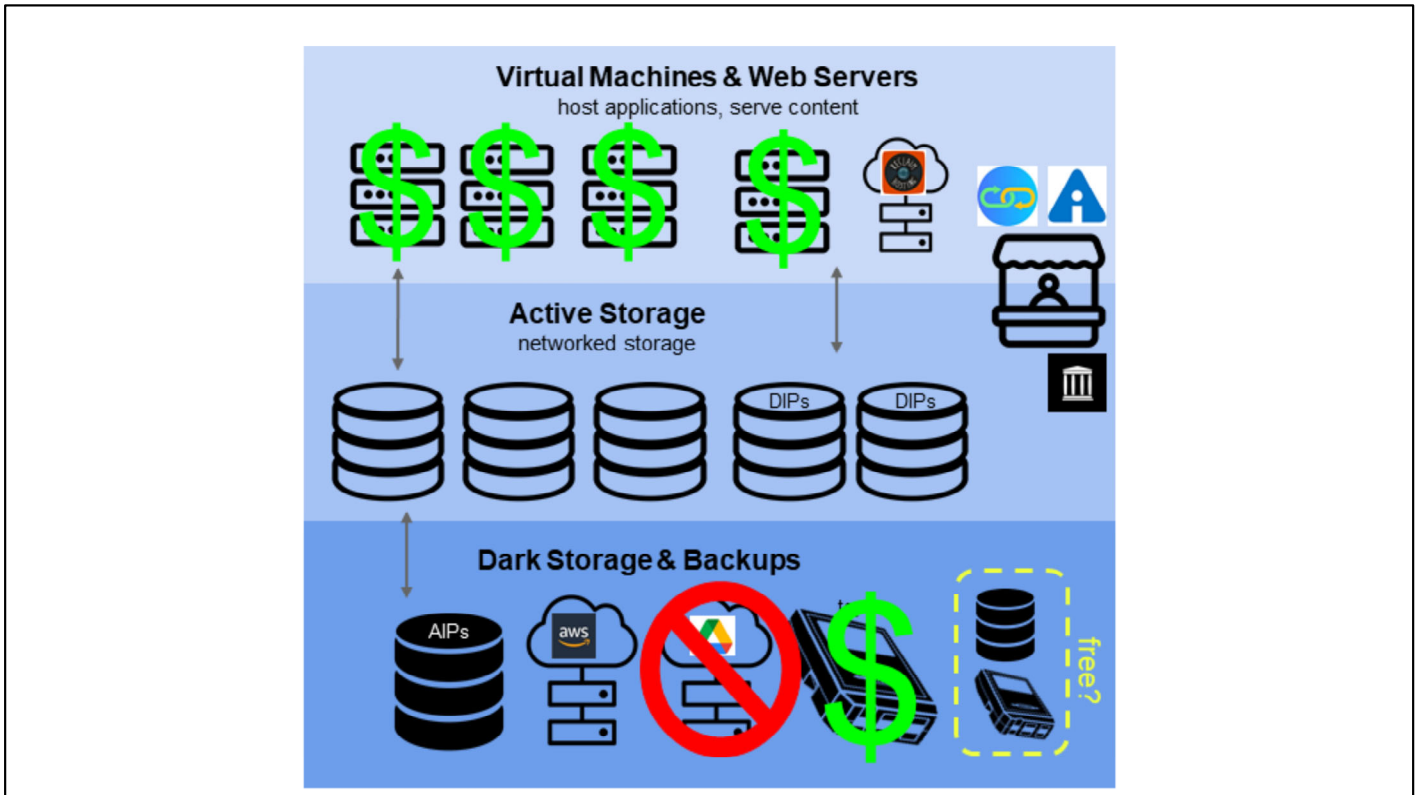
- **Change 1:** University ended its unlimited storage contract with Google, which essentially removed a backup location.



- **Change 2:** Started being charged for campus-supplied tape backups, which increased our annual costs by 35% (at current amount stored there)



- **Change 3:** Learned campus IT now offers “data protection services”:
  - Daily backups on 3 different types of storage media, data encryption, fixity testing
  - Meets the “sustain your content” level in the Storage category of NDSA’s levels of preservation
  - No charge



- **Change 4:** New pricing structure for UK ITS VMs and webservers (may double our annual infrastructure costs)

*These changes have had and will continue to have major implications for our resource allocation, workflows, and how we approach our digital curation work at UKL. We've responded to each of these changes individually to address the top concerns or opportunities, and have yet to do a comprehensive review and restructuring of our practices to account for the new variables.*

*That aside, I want to share some of our immediate responses.*



# Impact on Digital Curation and Preservation

	MINIMAL	INTERMEDIATE	FULL
<b>TRANSFER</b>	<ul style="list-style-type: none"> <li>☑️ Virus scan as needed*batch</li> <li>☑️ <a href="#">Substans</a> when appropriate</li> </ul>	<ul style="list-style-type: none"> <li>☑️ Virus scan as needed*</li> <li>☑️ <a href="#">Substans</a>, VLC, Windows Media Player, or Handbrake</li> </ul>	<ul style="list-style-type: none"> <li>☑️ Virus scan</li> <li>☑️ Disk image</li> <li>☑️ Storage</li> <li>☑️ File extraction- BitCurator Disk Image Access</li> </ul>
<b>STORE</b>	<ul style="list-style-type: none"> <li>☑️ Single standard</li> <li>☑️ Multiple backups</li> </ul>	<ul style="list-style-type: none"> <li>☑️ Single standard</li> <li>☑️ Multiple backups</li> </ul>	<ul style="list-style-type: none"> <li>☑️ Single standard</li> <li>☑️ Multiple backups</li> </ul>
<b>DOCUMENT</b>	<ul style="list-style-type: none"> <li>☑️ Update accession record and/or resource record</li> <li>☑️ Collection-level data in "born digital" spreadsheet (??)</li> </ul>	<ul style="list-style-type: none"> <li>☑️ Collection-level data in "born digital" spreadsheet</li> <li>☑️ Item-level media log**</li> <li>☑️ File info Exporter 3000 documentation file</li> <li>☑️ Update accession record and/or resource record</li> <li>☑️ ** PFI is especially visual or informative, capture information/network</li> </ul>	<ul style="list-style-type: none"> <li>☑️ Collection-level data in "born digital" spreadsheet</li> <li>☑️ Item-level media log</li> <li>☑️ Reports Bulk</li> <li>☑️ Curator Viewer and BitCurator Reporting Tool</li> <li>☑️ Update accession record</li> <li>☑️ Capture Digital PFI with description or visual info</li> </ul>
<b>REVIEW</b>		<ul style="list-style-type: none"> <li>☑️ Review files for challenging formats</li> </ul>	<ul style="list-style-type: none"> <li>☑️ Review files for challenging formats and file extension formats</li> </ul>

## Digital Preservation Policy University of Hull Library, Version 1.0 (2021)

### Introduction

The University of Hull Library has been assessing and creating digital resources since the 1980s. The digital content preservation and access to digital content continues to be central to UK Libraries and contributes to supporting excellence in research, education, and cultural engagement strategies. This policy formalises the existing commitment.

### Contextual Policies

UK Libraries Digital Preservation Policy should be approached in the context of other existing University, library and repository policies and strategies. These include the Library's existing University Strategic Plan and an agreed governance that may be developed in the future.

### Responsibilities & Mandates

UK Libraries mandate for the preservation of digital resources is published in missions and commitments at both the Library and the University level.

- **Accessioning** As a signpost for inclusion of digital resources, UKL is obliged to support accessioning, testing, and saving. To meet this need, an archive accession will be used to monitor in date and access arrangements. UK Libraries must also ensure that appropriate metadata is captured in accordance with the identification, description, and preservation of digital content.
- **Multi-tiered records** UK Libraries is responsible for maintaining records of access to digital resources and ensuring that records are available in digital, physical, or hybrid formats. This is a core part of the Library's mission.
- **Strategic priorities** UK Libraries Strategic Direction includes enhancing and strengthening access to research collections and supporting excellence throughout the research lifecycle. These will be strengthened by digital preservation practices that support the long-term retention of digital collections and ensure accessibility.
- **Contextual obligations** UK Libraries has contractual obligations to preserve and maintain access to certain digital resources. Some obligations extend from federal and state laws requiring the Library to report back for access reporting. This includes all

Decision tree

+

Digital preservation policy

[Link to UKL born-digital migration decision tree](#)

[Link to UKL digital preservation policy](#)

## Impact on Digital Curation and Preservation

We were fortunate to have two key documents to help respond and decide the most responsible path forward:

**Decision tree** to help make appraisal decisions for our born-digital resources. Megan will discuss this further.

**Digital preservation policy** (approved in July 2021) that

- Identifies content that is out of scope for digital preservation
- Digital assets that do qualify are divided into 4 tiers with increasing levels of preservation
- Articulates an institutional commitment to digital preservation (which is important with this quickly ballooning costs)





# Impact on Digital Curation and Preservation

	MINIMAL	INTERMEDIATE	FULL
<b>TRANSFER</b>	<ul style="list-style-type: none"> <li>☐ Virus scan as needed/batch</li> <li>☐ <a href="#">Substanz</a></li> <li>☐ when appropriate</li> </ul>	<ul style="list-style-type: none"> <li>☐ Virus scan as needed</li> <li>☐ <a href="#">Substanz</a>, VLC, Windows Media Player, or Handbrake</li> </ul>	<ul style="list-style-type: none"> <li>☑ Virus scan</li> <li>☑ Disk image</li> <li>☑ Storage</li> <li>☑ File extraction</li> <li>☑ BitCurator Disk Image Access</li> </ul>
<b>STORE</b>	<ul style="list-style-type: none"> <li>☐ Single standard</li> <li>☐ Multiple backups</li> </ul>	<ul style="list-style-type: none"> <li>☐ Single standard</li> <li>☐ Multiple backups</li> </ul>	<ul style="list-style-type: none"> <li>☑ Single standard</li> <li>☑ Multiple backups</li> </ul>
<b>DOCUMENT</b>	<ul style="list-style-type: none"> <li>☐ Update accession record and/or resource record</li> <li>☐ Collection-level data in "born digital" spreadsheet (??)</li> </ul>	<ul style="list-style-type: none"> <li>☐ Collection-level data in "born digital" spreadsheet</li> <li>☐ Item-level media log</li> <li>☐ File info Exporter</li> <li>☐ 3000 documentation file</li> <li>☐ Update accession record and/or resource record</li> <li>☐ If PII is especially visual or informative, capture information/work</li> </ul>	<ul style="list-style-type: none"> <li>☑ Collection-level data in "born digital" spreadsheet</li> <li>☑ Item-level media log</li> <li>☑ Reports Bulk</li> <li>☑ Custom viewer and BitCurator Reporting Tool</li> <li>☑ Update accession record</li> <li>☑ Capture Digital PII with description or visual info</li> </ul>
<b>REVIEW</b>		<ul style="list-style-type: none"> <li>☐ Review files for challenging formats</li> </ul>	<ul style="list-style-type: none"> <li>☑ Review files for challenging formats and file extension formats</li> </ul>

## Digital Preservation Policy

University of Manitoba Libraries, Version 1.0 (2021)

### Introduction

The University of Manitoba Libraries has been acquiring and creating digital resources since the 1960s. The long-term cultural preservation and access to digital content continues to be central to UoM Libraries and contributes to supporting excellence in research, education, and public engagement programs. This policy formalizes that continuing commitment.

### Contextual Policies

UoM Libraries Digital Preservation Policy should be approached in the context of other existing University, library, and repository policies and strategies. These include the Library's [Library Security and Access Policy](#) and its [Digital Preservation Policy](#).

### Responsibilities & Mandates

UoM Libraries' mandate for the preservation of digital resources is published in mission and vision statements at both the Library and the University level.

- Academic** As a signpost for inclusion of digital resources, UoM is engaged in research, teaching, and learning. The most accessible and diverse collection will ensure students in digital and other environments. UoM Libraries must therefore ensure that access to digital resources is supported by identification, description, and preservation of digital content.
- Infrastructure** UoM Libraries is responsible for maintaining currently active and ensuring that a secure and resilient digital infrastructure is maintained by the Library, Archives, and Records Commission of Manitoba.
- Strategic priorities** UoM Libraries Strategic Directions include enhancing and strengthening access to research collections and supporting excellence throughout the research lifecycle. These will be strengthened by digital preservation practices that support the long-term retention of electronic collections and ensure accessibility.
- Contextual obligations** UoM Libraries has a contextual obligation to preserve and maintain access to certain digital resources. Some obligations listed here, formal and otherwise, require the Library to respond to specific requests. This includes all

Decision tree

+

Digital preservation policy



*Thankfully, the sunsetting of unlimited Google Drive storage did not have a devastating impact on us-- we simply did a review of the content and cleared it.*



# Impact on Digital Curation and Preservation

	MINIMAL	INTERMEDIATE	FULL
<b>TRANSFER</b>	<ul style="list-style-type: none"> <li>☐ Virus scan as needed/batch</li> <li>☐ <a href="#">Substansys</a> when appropriate</li> </ul>	<ul style="list-style-type: none"> <li>☐ Virus scan as needed</li> <li>☐ <a href="#">Substansys</a>, VLC, Windows Media Player, or Handbrake</li> </ul>	<ul style="list-style-type: none"> <li>☑ Virus scan</li> <li>☑ Disk image</li> <li>☑ Compress</li> <li>☑ File extraction</li> <li>☑ BitCurator Disk Image Access</li> </ul>
<b>STORE</b>	<ul style="list-style-type: none"> <li>☐ Single standard backups</li> <li>☐ Multiple backups</li> </ul>	<ul style="list-style-type: none"> <li>☐ Single standard Multiple backups</li> </ul>	<ul style="list-style-type: none"> <li>☑ Single standard Multiple backups</li> </ul>
<b>DOCUMENT</b>	<ul style="list-style-type: none"> <li>☐ Update accession record and/or resource record</li> <li>☐ Collection-level data in "born digital" spreadsheet (??)</li> </ul>	<ul style="list-style-type: none"> <li>☐ Collection-level data in "born digital" spreadsheet</li> <li>☐ Item-level media log</li> <li>☐ File info Exporter 3000 documentation file</li> <li>☐ Update accession record and/or resource record</li> <li>☐ If PFI is especially visual or informative, capture information/work</li> </ul>	<ul style="list-style-type: none"> <li>☑ Collection-level data in "born digital" spreadsheet</li> <li>☑ Item-level media log</li> <li>☑ Reports Bulk</li> <li>☑ Custom viewer and BitCurator Reporting Tool</li> <li>☑ Update accession record</li> <li>☑ Capture Digital PFI with description or visual info</li> </ul>
<b>REVIEW</b>		<ul style="list-style-type: none"> <li>☐ Review files for challenging formats</li> </ul>	<ul style="list-style-type: none"> <li>☑ Review files for challenging formats and file extension formats</li> </ul>

## Digital Preservation Policy

University of Manitoba Libraries, Version 1.0 (2021)

### Introduction

The University of Manitoba Libraries has been acquiring and creating digital resources since the 1960s. The long-term cultural preservation and access to digital content continues to be central to UoM Libraries and contributes to supporting excellence in research, education, and public engagement programs. This policy formalizes that continuing commitment.

### Contextual Policies

UoM Libraries Digital Preservation Plans should be approached in the context of other existing University, library, and repository policies and strategies. These include the Library's overall University Strategic Plan as well as related policies that may be developed in the future.

### Responsibilities & Mandates

UoM Libraries' mandate for the preservation of digital resources is published in missions and commitments at both the Library and the University level:

- Academic:** As a signpost for inclusion of digital education, UoM is committed to support curriculum, teaching, and learning for those resources and services associated with those areas (research in digital and digital environments, UoM Libraries digital preservation activities must work in tandem to effectively support the identification, preservation, and presentation of integrated digital content.
- Institutional Records:** UoM Libraries is responsible for maintaining records created by teaching, learning and research departments, programs, research centres or divisions within UoM. This is a function of the records management services provided by the State Library, Archives, and Records Commission of Manitoba.
- Strategic priorities:** UoM Libraries Strategic Directions include enhancing and strengthening access to research collections and supporting excellence throughout the research lifecycle. These will be strengthened by digital preservation practices that support the long-term retention of electronic collections and create enduring scholarship.
- Contractual obligations:** UoM Libraries has contractual obligations to preserve and maintain access to certain digital resources. Some obligations have been formalized and may have specific contractual transfer fees for archive migration. This includes all

Decision tree

+

Digital preservation policy

## Dark Storage & Backups





*For the tape storage, we did some bulk purges according to our digital preservation policy that minimized the financial impact of this change.*



## Impact on Digital Curation and Preservation

Digital preservation cost estimate

1 TB = approx \$200 / year

(up from \$152 / year)



*Also, we now have updated cost estimates for preserving born digital content.*

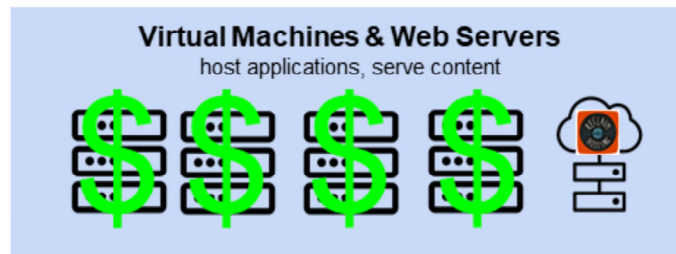
*1 TB = \$220/yr to preserve*



## Impact on Digital Curation and Preservation

### Funding challenges

How does this impact appraisal?



#### *Main questions---*

- *How do these changes (and essentially, new prices) impact our appraisal decisions?*
- *What capacity do we have for appropriate stewardship of collections?*

*The rest of the presentation provides examples of how these considerations have impacted our work*



## Takeaways

Be flexible and nimble.

*Build systems and workflows that are **flexible and nimble**. Many of these changes we experienced in the past 2 years were truly out of our control, but the diversity of our infrastructure allowed us to respond quickly and weather these changes.*



## Takeaways

Be flexible and nimble.

Build and maintain relationships with IT.

*It is important to build and maintain relationships with IT- Lib IT and Campus IT. They will know about potential infrastructure options and solutions than you may not be privy to.*



## Takeaways

Be flexible and nimble.

Build and maintain relationships with IT.

Storage costs increase over time.

*Drive it home to administration that digital preservation will only require more financial resources as time goes on. Collection acquisition rate outpaces any long-term decrease of storage costs. Robust infrastructure is expensive. Plan on 5-8% increase each year and adjust your budget accordingly*

All icons from the Noun Project <https://thenounproject.com/>







# Not everything is worth saving

Digital appraisal and environmental impact

Megan Mummey  
Director of Manuscript Collections

Andrew McDonnell  
Digital Archivist

Image from The Kentucky Quilt Project, inc. "Why Quilts Matter: History, Art, and Politics" documentary records

Hello again. Just a reminder that I'm Megan Mummey and for the next 10 minutes, I'll be presenting a case study on implementing earlier and more aggressive digital appraisal and the policy changes that resulted from this project. This version of the presentation is co-authored with my new Digital Archivist, Andrew McDonnell who couldn't be here today. Andrew has really taken the lead on the second half of this project. Recently, I have become more aware of the environmental impact of digital preservation. This forced me to come to terms with being "lazy" with my born digital archives. And maybe "lazy" is not the correct phrasing. "Under-resourced" and "under-staffed" are better ways to frame it. So I have always just "grabbed all the bits" and

said “This is future Megan’s problem”. Our original born digital archives program was very much focused on that “grabbing of all the bits” through disk imaging and wholesale data migration. The thinking was that we would figure out what to do with them later when they are processed (which realistically could be years down the road).

## Environmental Impact of Digital Preservation

“When [the challenges of digital preservation] are confronted in an environment where staff time is scarcer than digital storage, it can be tempting to appraise digital content in a cursory manner.”

Pendergrass, Keith L., Walker Sampson, Tim Walsh, and Laura Alagna. “Toward Environmentally Sustainable Digital Preservation.” *The American Archivist* 82, no. 1 (2019): 165–206.  
<https://www.jstor.org/stable/48659833>.

Before we dig into the case study, all of this work is rooted in the idea that Information Communication Technology (ICT) used by Cultural Heritage Organizations for things such as digital preservation and access has a negative environmental impact. I am not going to go into the arguments for why archivists should be aware of their environmental impact or why climate change is an issue or how big the impact of Cultural Heritage Organizations on the environment is. There is already plenty of scholarship on these issues. I am here today to share that reading this article on the screen and others like it made me aware that the decisions I make every day have an impact – and maybe there are things I can do to lessen that impact. I


have highlighted a quote that comes from the digital appraisal section of the article. “When [the challenges of born digital preservation] are confronted in an environment where staff time is scarcer than digital storage, it can be tempting to appraise digital content in a cursory manner.” That quote really resonates with me, and it makes me deeply uncomfortable, because this is what I have been directly doing. And I would just like to confront my privilege here – I’m at an R1 and I’ve never really worried about storage – though Sarah just outlined why I’m going to have to start worrying! I tend to do things the easy way and follow a procedure written by someone else and not critically examine what it is I am doing. I am having to retrain myself to confront the difficult tasks and decision points in my job – rather than kick that can down the road for someone else to deal with. This is all to say that perhaps we should be using the environmental impact of born digital records as a lens for asking ourselves the question – are these digital records in front of me worth saving?



HOME > EPISODES

## Episodes

*Why Quilts Matter: History, Art & Politics* is an exciting and thought-provoking nine-part documentary series that will change the way you think about quilts. Uncover the unsuspected size and depth of quilt culture. Delve into the inner workings of the art world, explore behind the scenes in the quilt marketplace, and journey into the complicated, kaleidoscopic world of quilts. Explore two centuries of history, watch as women gain power from their handiwork, and consider the nature of art itself.

 For an inside look at a highlight from each of the nine fascinating episodes of the documentary *Why Quilts Matter: History, Art & Politics*, please click on the episode listings below.



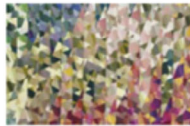
Episode 1: Quilts 101 –  
Antique and Contemporary



Episode 2: Quilts Bring  
History Alive



Episode 3: The Quilt  
Marketplace

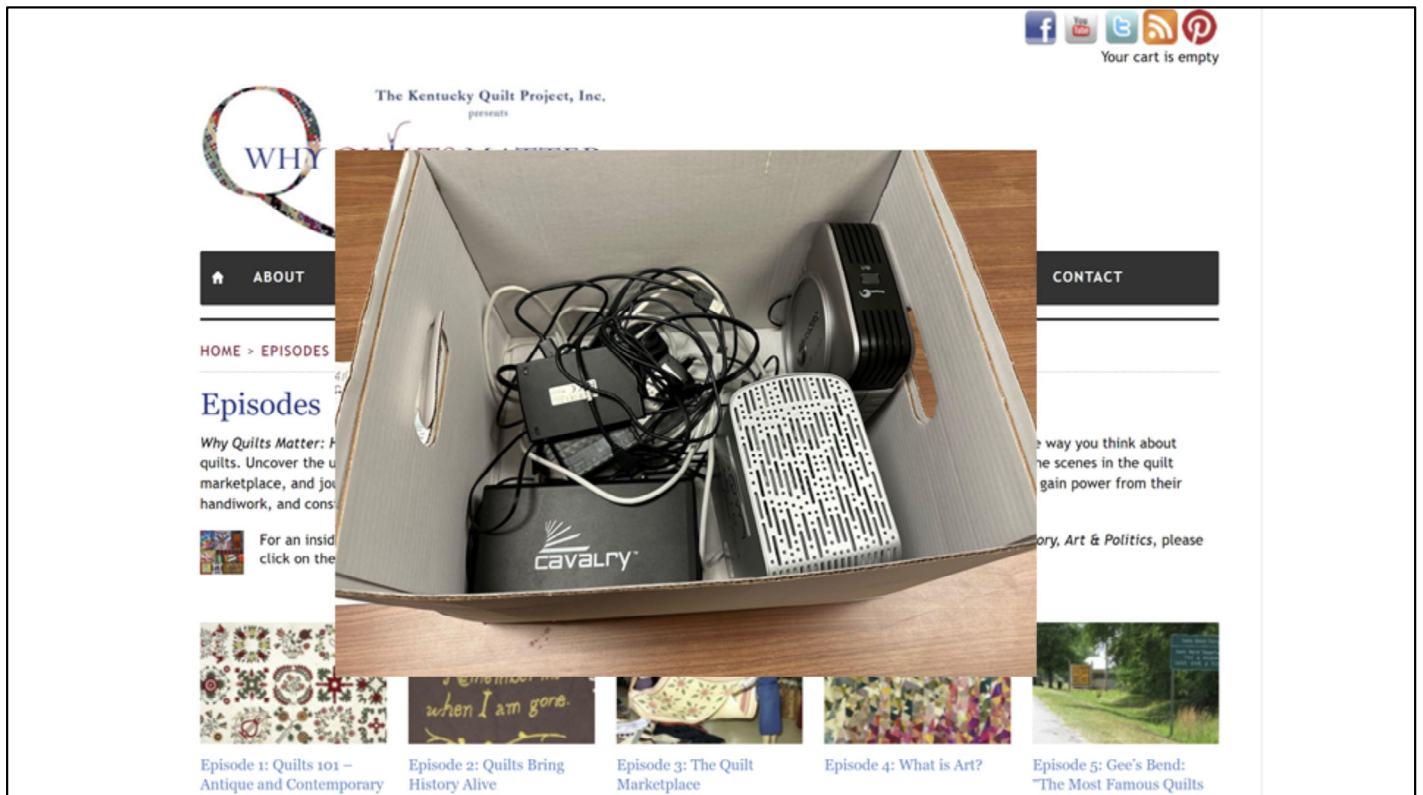


Episode 4: What is Art?



Episode 5: Gee's Bend:  
The Most Famous Quilts

In the Fall of 2022, myself and the director of the Nunn Center for Oral History, Doug Boyd – worked with a donor organization, the Kentucky Quilt Project, to bring in their records relating to a series of documentaries on quilts – *Why Quilts Matter*. The Nunn Center has a large collection of oral histories relating to quilting and this donation would bring in the original interviews done for the documentaries as well as the documentaries themselves. So everything I am going to talk about now has nothing to do with the management of the oral histories – those are being preserved and worked with very ably by my colleagues in the Nunn Center for Oral History.



We ended up with 4 record storage cartons, 6 hard drives, and 1 digital file transfer. The donor was very concerned that we get ALL the files, so she gave us everything she could find. The digital files added up over 3.5 TB. This is where normally I would have followed our documentation to a T and dutifully transferred the hard drives most likely using the minimal option on our migration decision tree, bagged everything, and backed it up in multiple places. However – sometimes following procedure to the letter – puts blinders on you. I am busy – we are all busy. And had I not paid attention to what I was preserving – I would have made a big mistake blindly preserving all of that data.

## Appraisal questions to ask

- What is the collection's archival value?
- Which records can we provide access to and approve the use of?
- Which records are essential to documenting this work/subject/person?
- Realistically – which records will patrons use?
- Do these files contain proprietary formats?
- Do we need to keep drafts of the documentaries?
- What at the bare minimum (given our capacity) should we preserve?

Here are the questions I asked myself during the appraisal of these records. And these are pretty normal appraisal criteria and questions. But they led me to the answer (using no tools) that yes some of these records have research value and fit with our collecting area. But there were major problems with many of the files foremost among them the proprietary formats and use issues.



Hard Drives/filetransfers	Size	No. files	Contents	Decision
HDD3 (seadisk)	58.5 GB	15,815	Image archive and image working files	Don't keep image archive except for photographs clearly provided by Shelly
HDD4 (new volume)	58.5 GB	15,816	Image archive and image working files	Don't keep image archive except for photographs clearly provided by Shelly
HDD2 (edit 12 backup)	1.76 TB	4,071	Image archive, final cut pro files	final cut pro files can't use - non destructive footage - proprietary can't use - old final cut pro only useful if you have old final cut
HDD1 (edit 13 backup)	1.59 TB	12,872	Cache files, DVD encodes, and "george" footage	Keep b-roll footage if any can be found mostly photographs - don't need cache files
file transfer	1.49 GB	4,448	Shelly's working files, website files, along with image archive	Don't need to keep website files
HDD5 (Seagate)	119 GB	83,149	Image archive	Don't keep image archive except for photographs clearly provided by Shelly
HDD6 (Seagate)	367 GB	55	Episodes 1-9 WQM, transferred from HDCAM Videotape masters, mp4 and MOV files	Keep - highest quality versions
<b>Received 6 Harddrives and 1 file transfer</b>	<b>3.5 TB</b>	<b>136,226</b>		

This spreadsheet shows you what was on each hard drive and in each file transfer. We got the original final cut pro files, the producers working files, website files, the cache files, the DVD encodes, their image archive (five different copies of it), as well as the high res copies of the documentaries themselves.

So out of these things – how much can we as an institution actually provide access to and grant use to? – the answer was very little.



Hard Drives/filetransfers	Size	No. files	Contents	Decision
HDD3 (seadisk)	58.5 GB	15,815	Image archive and image working files	Don't keep image archive except for photographs clearly provided by Shelly
HDD4 (new volume)	58.5 GB	15,816	Image archive and image working files	Don't keep image archive except for photographs clearly provided by Shelly
HDD2 (edit 12 backup)	1.76 TB	4,071	Image archive, final cut pro files	final cut pro files can't use - non destructive footage - proprietary can't use - old final cut pro only useful if you have old final cut
HDD1 (edit 13 backup)	1.59 TB	12,872	Cache files, DVD encodes, and "george" footage	Keep b-roll footage if any can be found mostly photographs - don't need cache files
file transfer	1.49 GB	4,448	Shelly's working files, website files, along with image archive	Don't need to keep website files
HDD5 (Seagate)	119 GB	83,149	Image archive	Don't keep image archive except for photographs clearly provided by Shelly
HDD6 (Seagate)	367 GB	55	Episodes 1-9 WQM, transferred from HDCAM Videotape masters, mp4 and MOV files	Keep - highest quality versions
<b>Received 6 Harddrives and 1 file transfer</b>	<b>368.5 GB</b>	<b>circa 4,000</b>		

We decided to not keep the image archive. Many of the images are in the one of the creator's archive at UofL and the rest of the images were licensed from other cultural heritage institutions and individual artists. And, no, that license did not include archiving the documentary records. We are keeping the lists of photographs considered which were generated by the documentary creators. As well as the image guides produced for each documentary – which list the image and the owning institution.

Next on the chopping block are the final cut pro files. These are not just files in a proprietary format, but files in a

2011 version of a proprietary format. Also – the reasons we are preserving this documentary have nothing to do with the making of the documentary.

Then we have cache files, dvd encodes, and footage from “George”. We also need none of these things. Originally, the “George” footage was identified by the donor as very important b-roll for the documentaries. And upon closer examination what was the b-roll? Another copy of that image archive. Nothing we can use. We will be keeping the master files for the documentaries and the producer’s working files (scripts and other planning documents) minus the image archive and the website files. This appraisal process first with minimal use of tools, through an environmental lens, gave us a good roadmap to start manipulating the files.

## Deduplication: TreeSize automated + manual

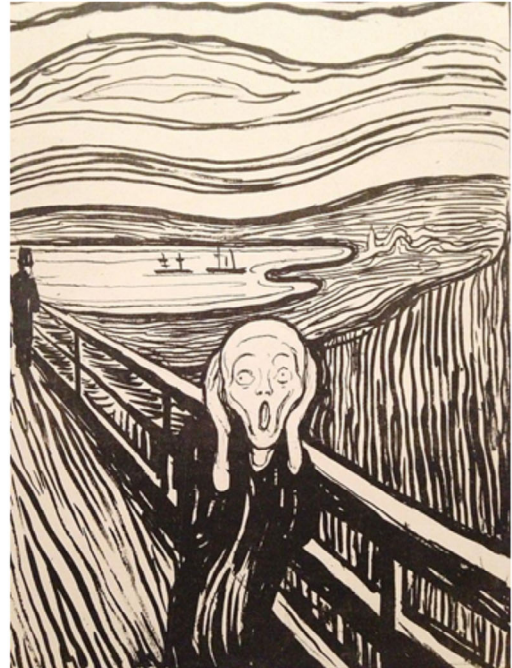
- ~60,000 duplicate files deleted (some recurred more than 10 times across multiple directories)
- Many duplicates were tiny, but in those quantities take up significant space and processing power
- Duplicate video files: multiple formats for DVD, broadcast, web, and other outlets
  - Manually discovered, fewer files but large file sizes!

Here is where I came in as the new Digital Archivist. I took this initial appraisal and started looking at the files with various tools, specifically TreeSize Professional and Bulk Extractor. After removing the Final Cut Pro file projects, containing multiple iterations of the documentary and cache files we didn't need, we used TreeSize Professional, a reasonably priced software tool, to further analyze the collection. Among its features, TreeSize can run a checksum analysis to find duplicate files, even when they have different file names. These automated searches allowed us to remove over 60,000 duplicate files (some of them recurred over 10 times across various directories). Many of them were tiny in size, but collectively added up to

over 200 GB. I also did a manual hunt for duplicate video files and found significant space savings, as videos were saved in Quicktime, WMV, mp4 formats for various outlets at different stages in the production and release process.

## PII and Tree-size analysis

Deleted folders and files of material irrelevant to the Kentucky Quilt Project, including personal vacation photos, medical files, legal files, real estate contracts, and other materials containing significant PII.



Another nice feature of TreeSize is it allows you to easily visualize what sorts of files are in your corpus, breaking things down to categories such as video, image, and text, but also allowing you to see where system files, empty directories, emails, and others are tucked away. In this case, the donor had full computer OS files in a large directory nested multiple directories deep, and TreeSize helped us track those down for deletion.

We found another significant space savings during PII analysis. I used Bulk Extractor to search for social security and credit card numbers, which on further exploration led to large collections of personal data that had found their

way into the collection, including family photos, medical files, personal email backups, real estate transactions, and other material that, in addition to being out of scope, created risk for the donor and for us as an institution.

## The Real Savings: Scratch Disk: 2.89 TB, 8,500 files



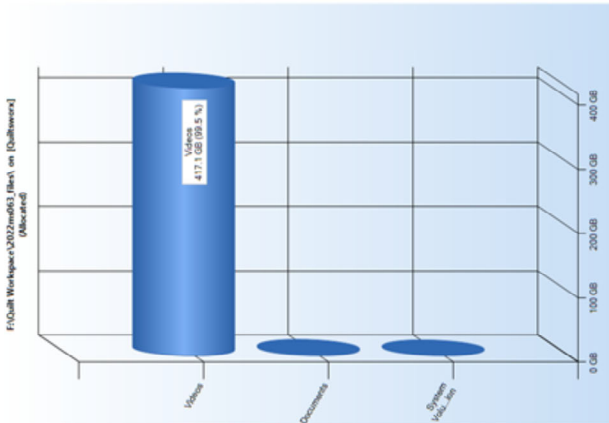
360 MB: 1:11 silent clip from a dark room

TreeSize allowed us to easily discover the vast, vast majority of disk space in the donation was composed of the scratch disks used in the video editing process. They included every single second of video shot over the course of the documentary: interviews, B-Roll, and even the videographer set-up or accidental shots, such as this minute-long clip of nothing but darkness and ambient noise, occupying 360 MB of space. In addition to saving our own storage, we were able to pass along the scratch disks to our oral history center, which had planned to spend thousands migrating interview footage from physical media for their collection.

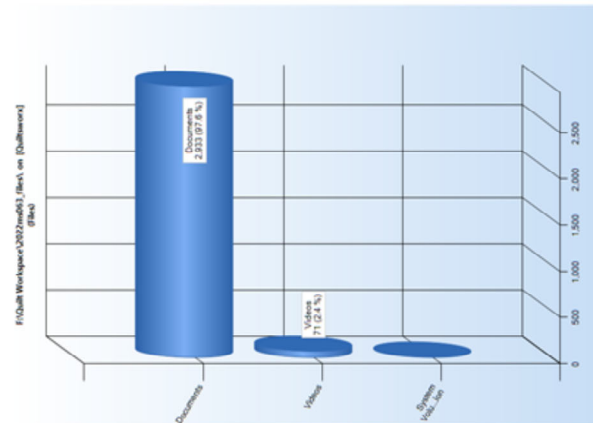
## Final Tally

- We started with: 146,664 files; 3.86 TB
- We conclude with: 3,006 files; 418.6 GB

- Cost for digital preservation: Original size= \$772/year; After Cleanup=\$83.60/year



Allocated space by file type



Number of Files

So we started with: 146,664 files; 3.86 TB. And ended with: 3,006 files; 418.6 GB. Which is 2% of the original file count and 10% of the original file size.

At an estimated cost of \$220/TB per year for digital preservation: Before Deaccessioning: \$849/year; After Cleanup: \$90.60/year

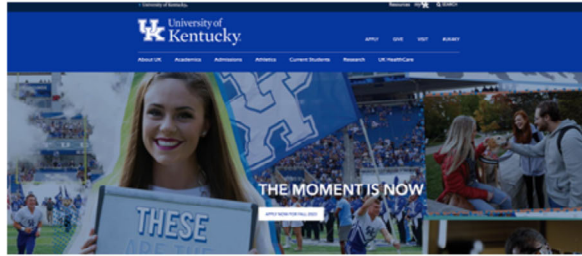


## Takeaways

- Added in a more rigorous appraisal section to our documentation.
- Saving everything is a trap. Do not do it! It's bad for archives and it's bad for the environment.
- Reappraise your born digital collections.

This whole experience has led us to re-envision our born digital processes with a sharp move away from the “grab all the bits” first approach. We’re working on a redo of our born digital workflows, all of this heavy lifting is being done by Andrew I might add. One of the main things that I want to make explicit in these workflows for those who engage in born digital work at UK, because there are quite a few of us, that they are encouraged to NOT SAVE EVERYTHING. Saving everything is a trap. It’s bad for archives and it’s bad for the environment. This experience has also encouraged me to embrace the concept of reappraisal – there are collections that I am now thinking about – that need to have their born digital records

reappraised, because we saved too much. Andrew is now in the process of evaluating collections that had previously been disk imaged. So to wrap up – Don't be like I used to be – don't kick the can down the road and just "grab everything" because it might be useful someday. In the end 98% of the digital records in this Why Quilt's Matter? Collection did not, in fact matter, and are not essential to preserve.



# Suddenly, Everything's Online! **What Do We Do Now?**

Ruth E. Bryan, CA  
University Archivist  
University of Kentucky Libraries  
Special Collections Research Center

We're going to pivot now to look at some of our challenges and strategies for online document preservation and description, and, since I'm Ruth Bryan, the University Archivist, this will be specifically for university records, to wit: "Suddenly, everything's online! What do we do now?"

## SUDDENLY! (The Realization)



=

2018

### Situation:

- Archive-It subscription
- Started crawling UK seeds
- Including yearly crawl of [www.uky.edu](http://www.uky.edu)
- Successful capture threshold = 75%

In 2018, Sarah and I were successful in advocating for the UK Libraries to purchase a subscription to Archive-It. We hired Emily as our first web archives intern and began selecting and crawling university websites and seeds. At that time, I thought of web archiving as just one of many acquisition and preservation methods for university records and allied documents. We also thought that we would be able to do appraisal, and set up crawls for all types of web content, including a yearly crawl for the main [uky.edu](http://uky.edu) seed, and after that one push of work, it would all be fine. We established that fine = 75% of the website or web page is captured, so our quality assurance threshold is good enough rather than perfect.

# SUDDENLY! (The Realization)



=

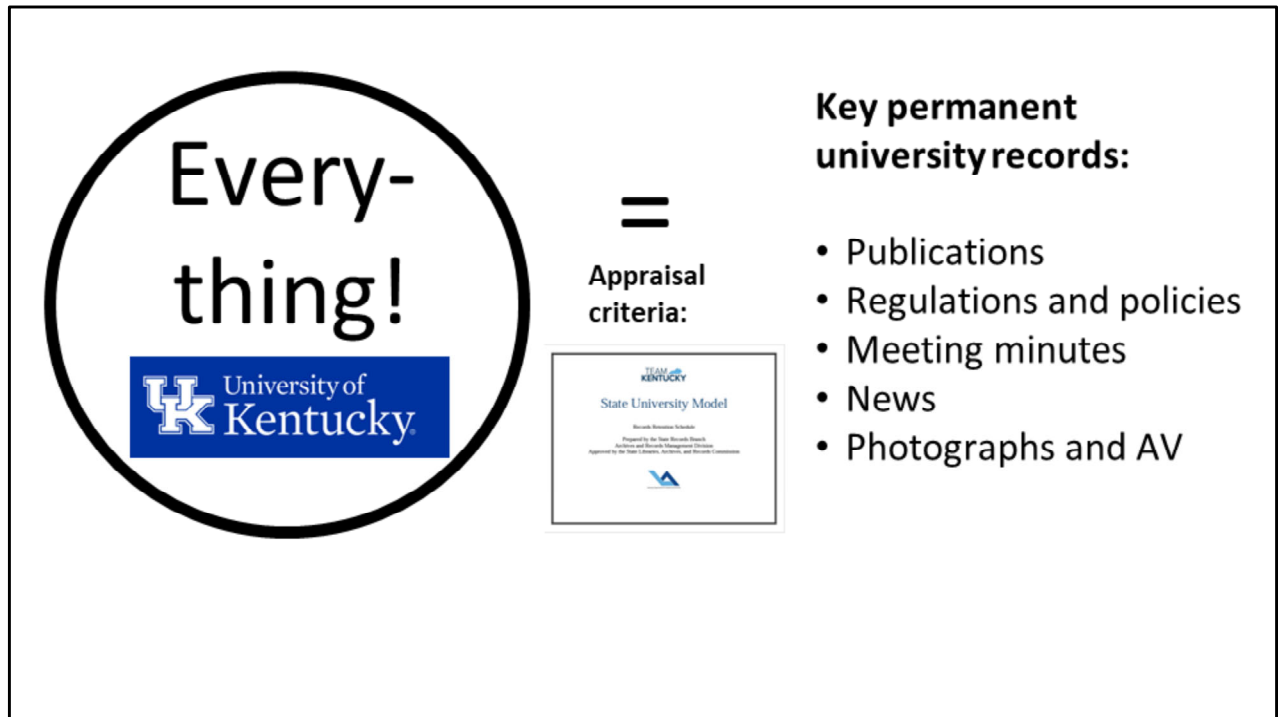
2019

## Situation:

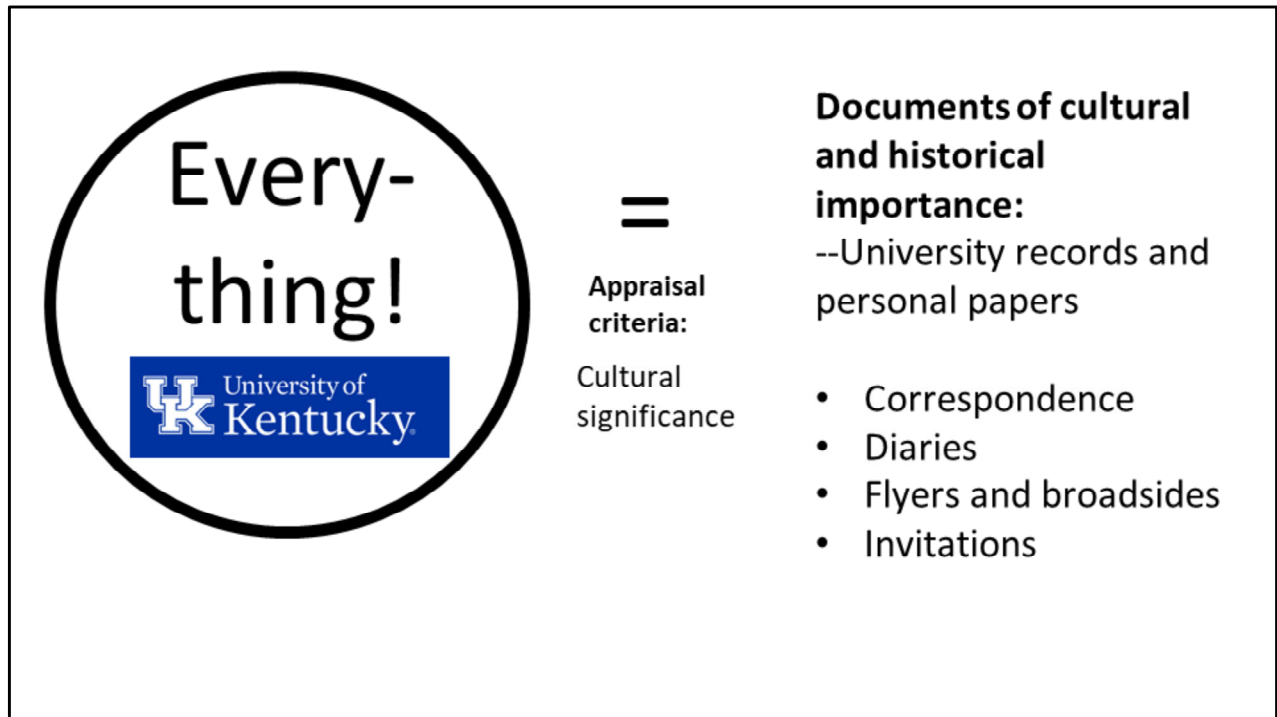
- Key records and papers
- Are distributed online only
- Not transferred to Archives
- Likely to be lost
- Web archiving is more central than before
- Preservation is complex
- Is thus resource-heavy
- Do we need to re-align effort?

But, by the next year, the realization washed over me that most key university records are being distributed or published online only and not routinely transferred to the archives the way they were in the past. The COVID pandemic accelerated this trend. So, web archiving is actually more important or more central than I had thought, because without proactively acquiring these documents, they are likely to be lost because of the ephemerality the web.

However, web archiving is technologically complex. It's not as easy as setting up the crawls, doing some quality assurance, and websites will be preserved forever. Web archiving requires more resources than many other formats. Do we need to re-align our effort?



The “Everything” in the presentation title refers to key, permanent University of Kentucky records as outlined in the State University Model Records Retention Schedule. These permanent records provide documentation of the university’s decisions and actions, finances, and planning.



“Everything” also includes other documents of cultural and historical importance that the records schedule considers non-permanent, but that are crucial to documenting the experiences and activities of university units and individuals. They often provide a counterweight to the official or public stance or story of the university.



Created by Ulcons

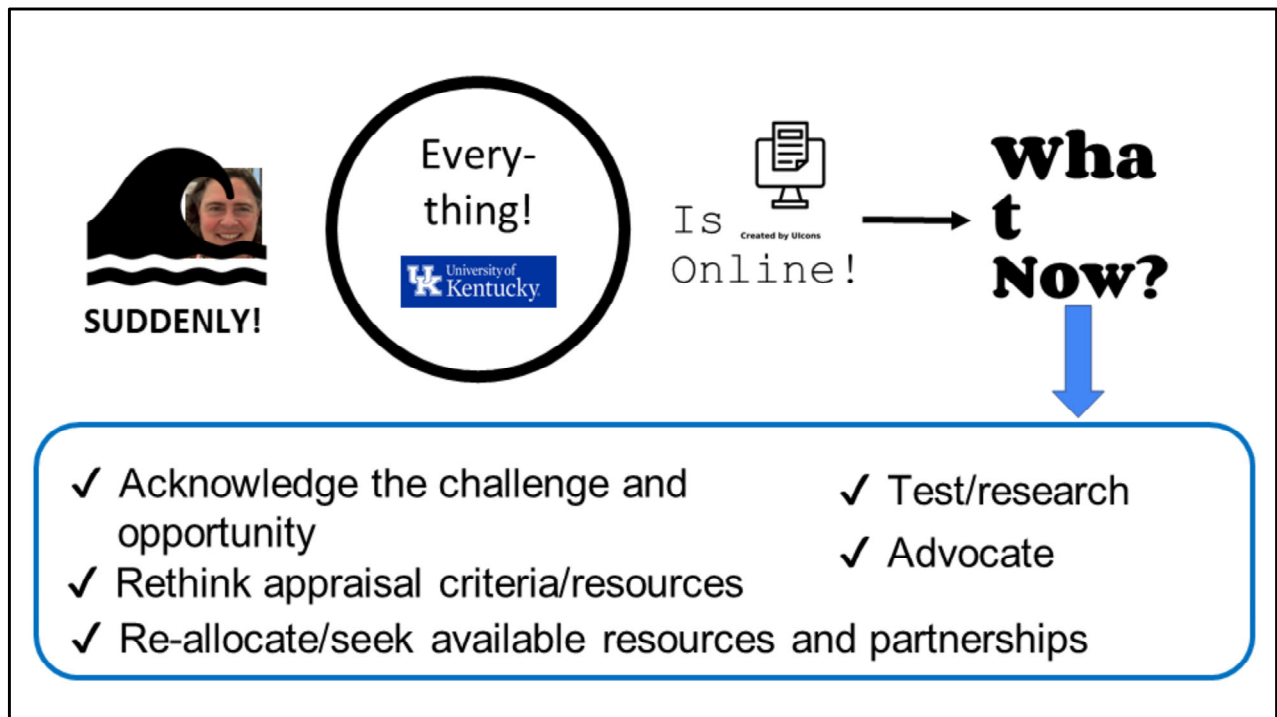
# Is Online!

- Many are in PDF
- Some PDFs on proprietary platforms with no download
- Websites/blogs
  - Embedded video not always captured.
  - Some webpages not captured at all (URL changes)
- Web 2.0 functions are difficult to capture
- Online ≠ Archived/preserved

The ways in which these key university records and other documents of historical value are distributed online varies considerably. Many are PDFs and can be downloaded or easily crawled. On the other hand, not many, but some important publications are distributed on proprietary platforms with no download option. Some are distributed as websites or blogs, but in an audit I conducted between 2021 and 2022, I found that some of the university's websites weren't captured at all. Social media dynamic scrolling, threading, commenting, and many other Web 2.0 functions are difficult to capture. And, records creators believe that putting documents online is the same as archiving them, so there's no need to send a copy to the archives. The archivist must now proactively search for and acquire these records.

online journal by Ulcons from [Noun Project](https://thenounproject.com/browse/icons/term/online-journal/ "online journal Icons")





To recap: I suddenly realized that key university documents are now being distributed online only. Managing these web-based documents is complex and requires additional resources. Given that I have scarce time and money, what do I do now?

First, I acknowledge the technological and resource challenge of online formats, plus the opportunity their acquisition provides for a wider, stronger presence of voices and content in the historical record.

Second, I rethink appraisal criteria, de-centering the university records schedule, prioritizing web-based documents created by underrepresented people and organizations, and more carefully quantifying the resources required for collection management.

Third, based on my appraisal, I re-allocate the resources I already have access to, and I seek out or respond to additional resources and relationships. Over the last few years, I have been able to request that most of my student budget be converted to continue to employ Emily as a part-time web archiving specialist, and now, as full-time Assistant University Archivist. This means stepping back from other formats, collections, and backlog projects for now. We were also awarded a mini-grant from

Project STAND to work with the Latino Student Union on their social media accounts, which is what Emily will be talking about. A neat thing that happened recently is that, because of our web archiving work, a web developer in the UK public relations office got in touch and is willing to help with preserving web sites! We just started this partnership.

Fourth, I continue to test and research to refine resource requirements and appraisal criteria.

Fifth, I use the research, testing, thinking, and practice to advocate for more support.

Even a small step means preserving key records, but your collection policy, institutional context, and existing resources will determine what “key records” are for you! Your future colleagues will thank you!

online journal by Ulcons from [Noun Project](https://thenounproject.com/browse/icons/term/online-journal/ "online journal Icons")

**Narrow Your Results**

Subject Sort By: Count | (A-Z)

- Hispanic American college students - Kentucky - Lexington - Societies, etc. (4)
- Student organizations - Kentucky - Lexington (4)
- Social media (3)
- Hispanic American college students-Kentucky-Lexington-Societies, etc. (1)
- Student organizations-Kentucky-Lexington (1)

Creator Sort By: Count | (A-Z)

- Latino Student Union (University of Kentucky) (4)

Type Sort By: Count | (A-Z)

- Online Social Network (3)
- Educational website (2)

Format Sort By: Count | (A-Z)

- Social Media (3)
- Website (2)

Coverage Sort By: Count | (A-Z)

- Kentucky (5)
- Lexington, Ky. (5)

Collector Sort By: Count | (A-Z)

- University of Kentucky Libraries (5)
- Special Collections Research Center (4)
- University of Kentucky Libraries Special Collections Research Center (3)

Identifier Sort By: Count | (A-Z)

- 2019su043 (5)

Sites for this collection are listed below. Narrow your results at left, or enter a search query below to find a site, specific URL, or to search the text of archived webpages.

Enter search terms here

Group: Latino Student Union

Sites

Page 1 of 1 (5 Total Results)

Sort By: Title (A-Z) | Title (Z-A) | URL (A-Z) | URL (Z-A)

**Title:** Twitter. Latino Student Union (University of Kentucky) (@lsu\_uky) social media account

**URL:** [https://twitter.com/lsu\\_uky/](https://twitter.com/lsu_uky/)

Captured 27 times between September 21, 2019 and May 20, 2023

**Video:** 14 Videos Captured

**Subject:** Student organizations - Kentucky - Lexington. Social media. Hispanic American college students - Kentucky - Lexington - Societies, etc.

**Group:** Latino Student Union

**Creator:** Latino Student Union (University of Kentucky)

**Coverage:** Lexington, Ky., Kentucky

**Format:** Social Media

**Type:** Online Social Network

**Identifier:** 2019su043

**Related Materials:** Forms part of the Latino Student Union (University of Kentucky) records, accession number 2019su043, located in the Special Collections Research Center, University of Kentucky Libraries.

**Collector:** University of Kentucky Libraries. Special Collections Research Center

---

**Title:** Wixsite. Latino Student Union (University of Kentucky)

**URL:** <https://uklatinouni.wixsite.com/lsuuky/>

Captured 11 times between September 21, 2019 and May 20, 2023

**Subject:** Hispanic American college students - Kentucky - Lexington - Societies, etc., Student organizations - Kentucky - Lexington

**Group:** Latino Student Union

**Creator:** Latino Student Union (University of Kentucky)

**Coverage:** Lexington, Ky., Kentucky

**Format:** Website

**Type:** Educational website

**Identifier:** 2019su043

**Related Materials:** Forms part of the Latino Student Union (University of Kentucky) records, accession number 2019su043, located in the Special Collections Research Center, University of Kentucky Libraries.

**Collector:** Special Collections Research Center, University of Kentucky Libraries

## Wildcat Histories: Preserving Activist UK Student Organization's Legacies

Emily Collier  
 Assistant University Archivist  
 Web Archiving Specialist  
 Special Collections Research Center  
 University of Kentucky

Seeds belonging to the Latino Student Union  
 on UK Libraries' public Archive-It page

Hello, I am Emily Collier, the Assistant University Archivist and Web Archiving Specialist for Special Collections, and I will be diving into our partnership with the Latino Student Union and our efforts to archive the cultural heritage found in social media, a much more tricky type of online resource to capture and preserve.



## The page cannot be found

The page you are looking for might have been removed, had its name changed, or is temporarily unavailable.

---

Please try the following:

- If you typed the page address in the Address bar, make sure that it is spelled correctly.
- Click the [↶ Back](#) button to try another link.
- Click [🔍 Search](#) to look for information on the Internet.

HTTP 404 - File not found  
Internet Explorer

## Key Problems:

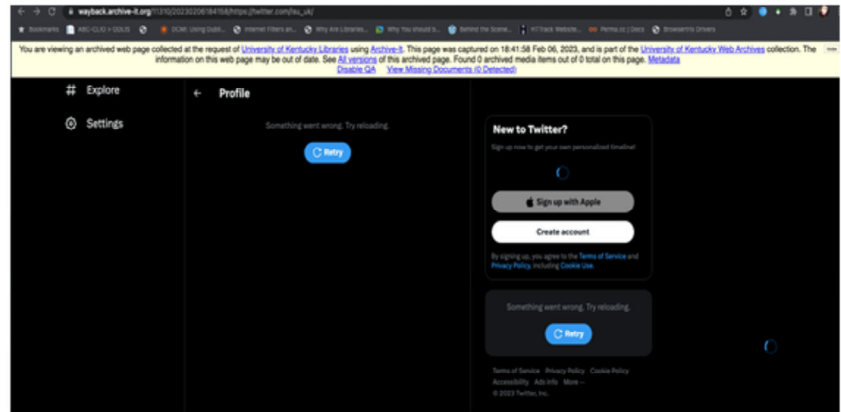
- Link rot
- Archiving tool failures
- Website upgrades
- Enterprise interception

So some of us who have worked with web archiving know many of the key problems when facing any website, including link rot, archiving tool failure, upgrades to sites like the inclusion of dynamic or interactive content, and of course when website proprietors make changes to their sites or install things like permissions. Sites can be changed, moved, taken down, become hidden behind paywalls or logins, and couple this with the struggle of web archiving tools to capture dynamic scripts, and you can really end up in the weeds.

# Social Media: Our Notorious Evader

## WHY?

- Complex, interactive scripts
- Crawler traps/infinite links
- Near-constant updates create an arms race with preservation technology
- Crawler blockers and permissions obstacles



A failed Twitter crawl with Archive-It

And as you can expect with those complex and interactive scripts, social media is the absolute worst to try and capture. Social media also often contains crawler traps, such as infinite links, meaning we have to be more careful about our scoping practices or else we end up with a ton of stuff we don't want. They are also more subject to enterprise interception. These sites go through constant updates and changes to formatting and it really is an arms race for many web archiving tool developers to ensure their tools are able to work through those updates. What's more, many of these sites specifically have crawler blockers included in their scripts and also prevent content from being viewed without being logged in. So here you can see a failed Twitter crawl from Archive-It. Archive-It has had quite a hard time with most social media so one of the exciting parts of our student organization partnership is the opportunity to research other options.

## Wildcat Histories: A Flagship

- A Project STAND (**ST**udent **A**ctivism **N**ow **D**ocumented) mini-grant
- IMLS and Mellon Foundation funding
- April 2022-Aug 2023
- Partnered with the Latino Student Union
- Use the partnership as a pilot for building procedures on archiving student organizations' online content, specifically social media



So here we are with our Wildcat Histories project. Project STAND has been around for about 5 years and attracted Ruth's attention due to its focus on ethical documentation of student activism in marginalized or underrepresented communities. As she already had a working relationship with the Latino Student Union and a small collection of their materials, she approached them for a partnership. We received a \$14,000 grant through Project STAND that was funded through IMLS and the Mellon Foundation for work to be completed from April 2022 through August of 2023. The goal was to use Wildcat Histories as a pilot for building successful procedures for archiving student organization's online content, specifically social media.

# Social Media Goals

## Theoretical Goals



### Preserving Memory

Appraising platforms that appropriately capture communities' memories



### Social Interactions

Ensuring that the social interactions, ie comment sections, are preserved



### Online Culture

Preserve the unique qualities of online culture

## Practical Goals



### Testing

Test available tools and methods for best capture and preservation practices



### Documentation

Document tests and outcomes in order to keep a record of chosen methods and rationale



### Standardization

Use these rationals to create a standardization of procedures and application of metadata

So here I have outlined the project goals for my part. The theoretical goals include Preserving memory, right. Understanding and appraising those social media sites to ensure we are appropriately capturing the voices of the LSU members. It's easy to grab extra content when web archiving so this first step ensures I am capturing the voices without grabbing links that go too far out of context. It is also important for us to preserve social interactions on these sites, so the comment section being a really great place for this. It gives us context. It gives us a better understanding of the opinions of community members and it also allows us to see trends in thoughts. All of this culminates into preserving the online culture of a group, which is distinct. AND it really gives us a unique perspective into the functions of a group. Social media allows for pictures and videos and conversations that you just don't get with printed meeting minutes or flyers. This leads me to our practical goals. My role is to test the current tools available in order to find the ones that work the best for capturing those theoretical goals. I'm also keeping documentation of these tests and their outcomes as this gives me a chance to decide which methods and practices are the best. The rationals I make then go into developing a standardization of procedures and best practices for capturing social media sites.

## Student Archiving Goals

- Empowerment: Return control to student organizations
- Community: Build positive and productive relationships between students and Special Collections
- Sustainability: Create solutions to integrate archiving into student organization policy

**WILDCAT HISTORIES**

This project focuses on the preservation of student organization social media content in University Archives. Organizations for members identifying as BIPOC, LGBTQ+, international, not Christian, and/or women in male-dominated disciplines are a key part of student life, building support networks, having fun and serving their communities, and raising and acting on concerns and challenges.

**WHO?**

Organizations for members identifying as BIPOC, LGBTQ+, international, not Christian, and/or women in male-dominated disciplines.

The primary aim of **Wildcat Histories** is to help you document and preserve your own social media. This lets you preserve your legacy and retain important organizational information that would otherwise be lost in the digital realm.

**WHY?**

A way to sustain your organization's legacy, provide continuity, conservation, and organization

Follow [the link](#) to see an example of how your organization can be highlighted by University Archives!

**CONTACT US**

taylor.c.leigh@uky.edu  
nuth.bryson@uky.edu  
claudiaolizbeth.benito@uky.edu

UK Libraries

Informational flyer created by our Student Organizations Communication Assistant, Claudia Benito

We also have another goal, and that is to get the student groups themselves invested in their own archives. This is important because it gives control of preserved content to the organizations themselves allowing them to pick and choose what they want preserved. It builds a positive and productive relationship between these groups and Special Collections, and it also increases the chance of sustainability with more hands on-deck. If we can find the easiest methods and tools to use for archiving, match that with a successful delivery of the value of using archives to support an organization's legacy, the more likely the students will be take on their own archiving practices.



## Building Forward-Thinking Practice Around Anticipating Failures

Gaps can be left in collections when crawling technologies fail.

Anticipating these failures and preparing backup tools or policy can alleviate this burden.



The same Twitter page, successfully captured with Webrecorder

But we need to find those easy solutions first. As I mentioned before, web archiving tools frequently fail with social media, and as tool developers work to keep up with these changes, the down time can result in gaps in our collections. So we want to build a more forward-thinking practice by expecting failures. By anticipating gaps and thinking about how our policy and practice can adapt, we can better prepare ourselves for failures. We are doing this currently by researching more web archiving tools and building a sort of backup arsenal to turn to when Archive-It fails. So Webrecorder, specifically their archiveweb.page plugin, has so far been quite successful where Archive-it fails. Earlier, I showed a slide with a failed Twitter crawl with Archive-It, and here is that same page I managed to get with Webrecorder a couple weeks later. I'd like to stress here that Webrecorder was a tool we had looked in 2019, but had dismissed at the time because we weren't super impressed with it.

## Redundancy in Research

- New technologies always coming into the field
- Existing technologies typically successful may fail in the future
- Old technologies once dismissed may have upgraded to better tools



And this brings me to one of the great take aways of the Wildcat Histories project, which was to learn about the value of redundant research. Web archiving is extremely volatile work. New technologies are always coming into the field, but almost just as importantly, old technology we once dismissed may evolve into much better tools. You can't just look at a tool and totally trash it if it doesn't seem like it will work for you because down the line, it may turn out to be just what you need. Especially when your existing tools start to fail.

### Questions Proposed

# Q

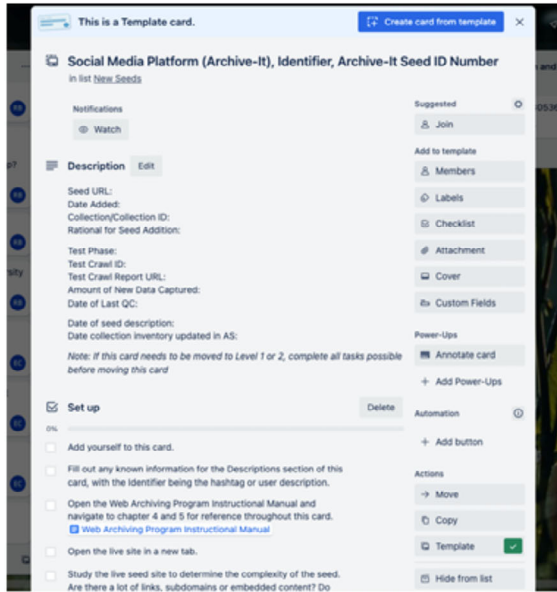
1. What are your tools, how do they work, and what are their products?
2. How is content made accessible to users?
3. How is your program funded/what is your long term management plan?
4. What are your next steps/developments?

### Professional Network

# A

- **Peter Chan**, Stanford University Libraries
- **Zakiya Collier**, Documenting the Now
- **Lynda Schmitz Fuhrig**, Smithsonian Libraries and Archives
- **Bergis Jules**, Documenting the Now
- **Christie Moffatt**, National Library of Medicine
- **Jasmine Mulliken**, Stanford University Press
- **Dolsy Smith**, Social Feed Manager
- **Ed Sommers**, Documenting the Now, Stanford University Libraries
- **Brian Thomas**, Texas State Archives

But we do have to be careful about what tools we adopt. My position only allots a certain amount of time I can dedicate to web archiving. I don't have the time to invest in tools that won't work for us. So rather than take hours to investigate and test new tools, we have used the Wildcat Histories project as a chance to reach out to a professional network of archivists and developers to ask them about their experiences. What tools have they created or used that worked well for them? What kind of outputs do their tools create and how are those outputs made compatible with their existing collections? How is the content in those collections made accessible to users? If we speak with developers, we also want to know how their tool is being funded for maintenance, support and development. The sort of elephant in the room about open source tools is that they can be great solutions, but they can also very risky and not cost effective even if they're "free". If the main developer retires, will that tool go away? Will someone take over? What if the tool was grant funded and the money runs out? If I've invested my time in a tool, I've invested money. If the tool fails and I have to do an overhaul, good chance I've wasted some money. That being said, I'd like to put in here that if an open source tool is truly valuable, the benefits may outweigh the risks and it is important to support these tools so they can be further developed and maintained. It really is about finding a balance and what works best for your institutional needs.



Trello card with steps for crawling social media sites with archive-it

## Maintaining Simplicity

- Adopting complex technologies or practices makes long-term management difficult
- If a staff member leaves and takes knowledge with them, how is the work maintained?
- Needs to be simple for students to understand and use readily

For us, one of our institutional needs is maintaining simplicity. If we adopt too many tools, it's going to be too burdensome to maintain workable practices and documentation with our limited resources. And because we do have limited resources, another thing we have to keep in mind is the skill-level of technology we can properly maintain. Will we always have a staff member that can use complex tools? If that staff member leaves and takes all their knowledge with them, how do we maintain the work? The simpler and more supported the tools are that I adopt into our practices, the more sustainable. I have also crafted our technology workflows around a REALLY thorough set of procedures and VERY descriptive (and pictured!) instructions in order to enhance longevity of our web archives. Basically my goal is that almost any new staff member should be able to read my documentation and complete the very base work satisfactorily, even students. And this has been a major part of Wildcat Histories. We want students to take an active role in archiving their own content. Web archiving can be difficult, particularly with social media, but if we can determine the right tools and if the methodology is simple enough, even students can understand it and use it readily. Special Collections will always be here to problem-solve and provide updates to changes, but we want the students to feel confident about using web archiving technology and empower them to take part in preserving their own legacies.

## Themes

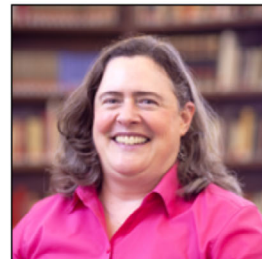
- Documentation is important
- Flexibility is key
- Relationship building can open doors
- Digital preservation requires funding and personnel
- Record procedure or policy changes, *with* rationales
- View workflows as learning-in-working or practiced-based and adaptable, rather than a set of rigid tasks
- Short-term and ongoing re-appraisal and re-prioritization and re-allocation of existing funds

## Future directions and questions

- Reappraisal of born digital materials already migrated and preserved, but not processed
- Access tools for web archives, including data harvesting tools
- Resource allocation and workflows for description and capturing are different for web-based content vs. other born-digital content
- Is appraisal criteria for permanent and culturally significant online university records different from other born-digital content?
- Should we apply digital preservation criteria retroactively? This requires additional development for our tools

## Q&A

- Sarah Dorpinghaus, Director of Digital Strategies and Technology
- Megan Mummey, Director of Manuscript Collections
- Andrew McDonnell, Digital Archivist
- Ruth E. Bryan, CA, University Archivist
- Emily B. Collier, Assistant University Archivist



You will be hearing from us in the following order – First you will hear from Sarah Dorpinghaus about shifting digital preservation infrastructure, then myself on implementing born digital appraisal, then Ruth Bryan on the acquisition of university publications, and then Emily Collier on web preservation. It may seem like we are all talking about disparate subjects, but each presentation will build on each other to form an in-depth case study of how we have been attempting to wrangle the beast that is working with born digital materials.

So if you know me – you know I have a tendency to say flip things (because I'm a youngest child so I'm always

trying to get a laugh). I often say things like “that’s future Megan’s problem”. But I’ve been an archivist for enough time now that I when run across problems, I get angry, and say “who did this!?”...and it’s always “past Megan”. So this panel came together upon the realization that we are all trying to not do this. We are struggling with various pain points, like time, expertise, understaffing, and trying our best to plan for the future in the constantly changing landscape around digital stewardship.