6-11-2024

This Be the Beloved Curse: Learning to Love Ever-Evolving Born-Digital Description

Ruth E. Bryan
University of Kentucky Libraries, ruth.bryan@uky.edu

Megan M. Mummey
University of Kentucky Libraries, megan.mummey@uky.edu

Andrew McDonnell
University of Kentucky Libraries, mcdonnell@uky.edu

This Be the Beloved Curse:

Learning to Love Ever-Evolving Born-Digital Description

Ruth E. Bryan, Andrew McDonnell, Megan Mummey
Session 13
June 12, 2024

The title of the presentation, “This be the beloved curse” is taken from Philip Larkin’s poem “This be the curse.” In it, Larkin describes the cycle of life where children are messed up by their parents, who, with the best intentions, “...fill you with the faults they had/and add some extra, just for you.” He goes on to explain, though, that our parents, in their turn, were equally messed up “...by fools in old-style hats and coats…”, handing on what Larkin terms “misery” from one generation to another.

As archivists, and especially as archivists working with digital formats, our “beloved curse” is the constant change of technology and archival praxis. We make the best or at least the most practical collection management decisions possible today for preservation and access based on repository mission and the resources and expertise available at the time. We mean well, but the decisions we make today are doomed to obsolescence. To our successors, our choices will seem uniformed, ineffective, or even bizarre. For our part, we, too, believe our predecessors were, in the main, crazy. “Why did they do it this way?!?” we moan.

In the final line of the poem, Larkin advises his readers to not “...have any kids yourself,” but this isn’t an option for archivists. Our collections live forever! In order to continue to be effective stewards of digital-format documents, archivists must routinely revisit the collection management decisions of predecessors at their repository and do our best for our collections, trying to pass on as little “misery” as possible.

Outline

- Case Study 1: Kentucky League of Cities records
  - Megan Mummey, Director of Manuscript Collections
  - Andrew McDonnell, Digital Archivist
- Case Study 2: UK Athletics Film and Video collection
  - Ruth E. Bryan, CA, University Archivist
  - Andrew McDonnell, Digital Archivist


Megan (she/her)


Andrew (he/him)


Ruth (she/her)

In this presentation, we'll share two case studies demonstrating how, at the University of Kentucky Libraries Special Collections Research Center and, with the hiring last year of a new Digital Archivist, we're learning to embrace our beloved curse of constant change, focusing on the area of arrangement and description. The case studies will demonstrate how previous archivists described digital-format personal papers and university records, asses those approaches, and share new decision trees, tools, and procedures we are creating and implementing now. Megan Mummey (she/her pronouns), Director of Manuscript Collections, and Andrew McDonnell (he/him pronouns), Digital Archivist, will lead off with our first case study about born-digital description in manuscript collections centered on the Kentucky League of Cities records. Ruth Bryan (she/her pronouns), University Archivist, and Andrew will continue with a second case study from the University of Kentucky Athletics Film and Video collection. We hope to demonstrate that, rather than cursing our predecessors (and ourselves), we're learning to love ever-evolving description of digital formats as a natural part of the digital lifecycle. Andrew can't join us in person today, so he recorded his portions of the presentation.

UK Institutional Context

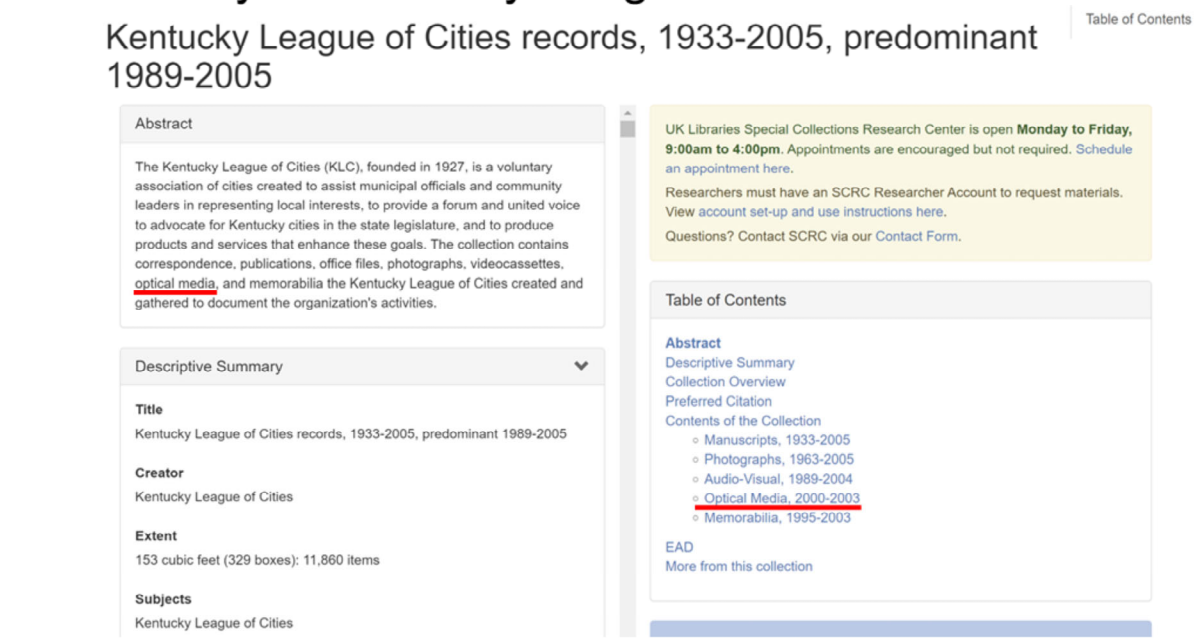
- UK is a research 1, land-grant institution in Commonwealth of Kentucky
- UK Libraries' digital preservation repository and digital library are both home-grown systems
- UK Libraries has engaged in formal born-digital work since 2015
- 1 FTE 100% born digital archiving, 9 FTE partial (5-15%), and on average 0-3 student employees
- Born digital archiving and digitization are a part of the SCRC

Margaret I. King Library

Now let us dive into our first born digital description case study! The Kentucky League of Cities records. They are A MESS. But a mess created with good intentions. The finding aid is the earliest example of an attempt to describe born digital records for public consumption at UKL, that I could find. The processing began upon acquisition in 2005 and was completed in 2007. This was also before any of us came to the UKL. So, we have no institutional knowledge! Hurray! I also want to take a moment to say that for 2005 this was really forward thinking for its time! They were trying to do something to provide access to these files instead of letting them sit there undescribed! So, kudos but it is time for this finding aid to get a glow-up! In 2017, we had a graduate student focus on migrating files from physical media to our servers. They disk imaged and migrated the Kentucky League of Cities digital files, amounting to 4.4 GB and almost 7000 digital files. However, the

project only focused on migration and not description or access. So, there is a lot of work to still be done.

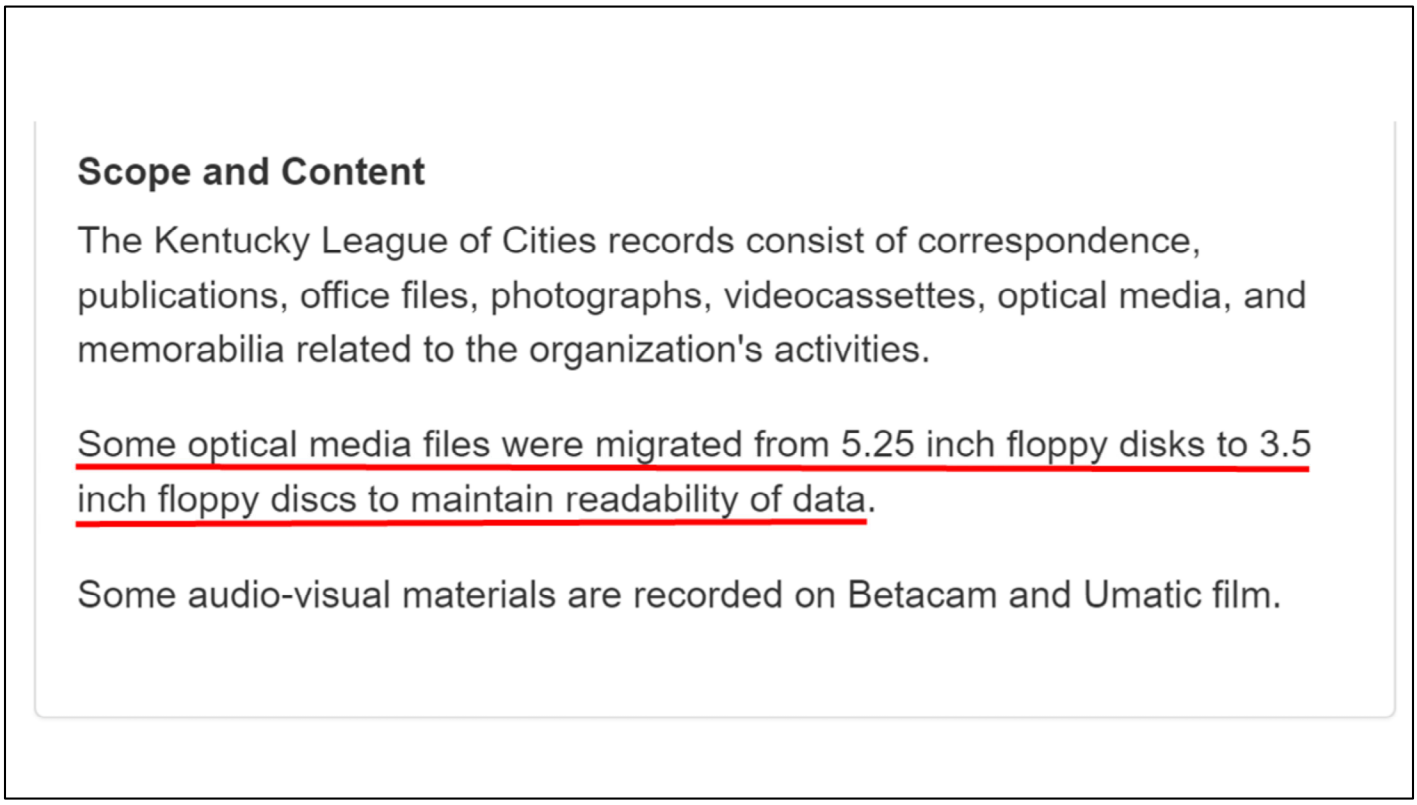
Case Study 1: Kentucky League of Cities records



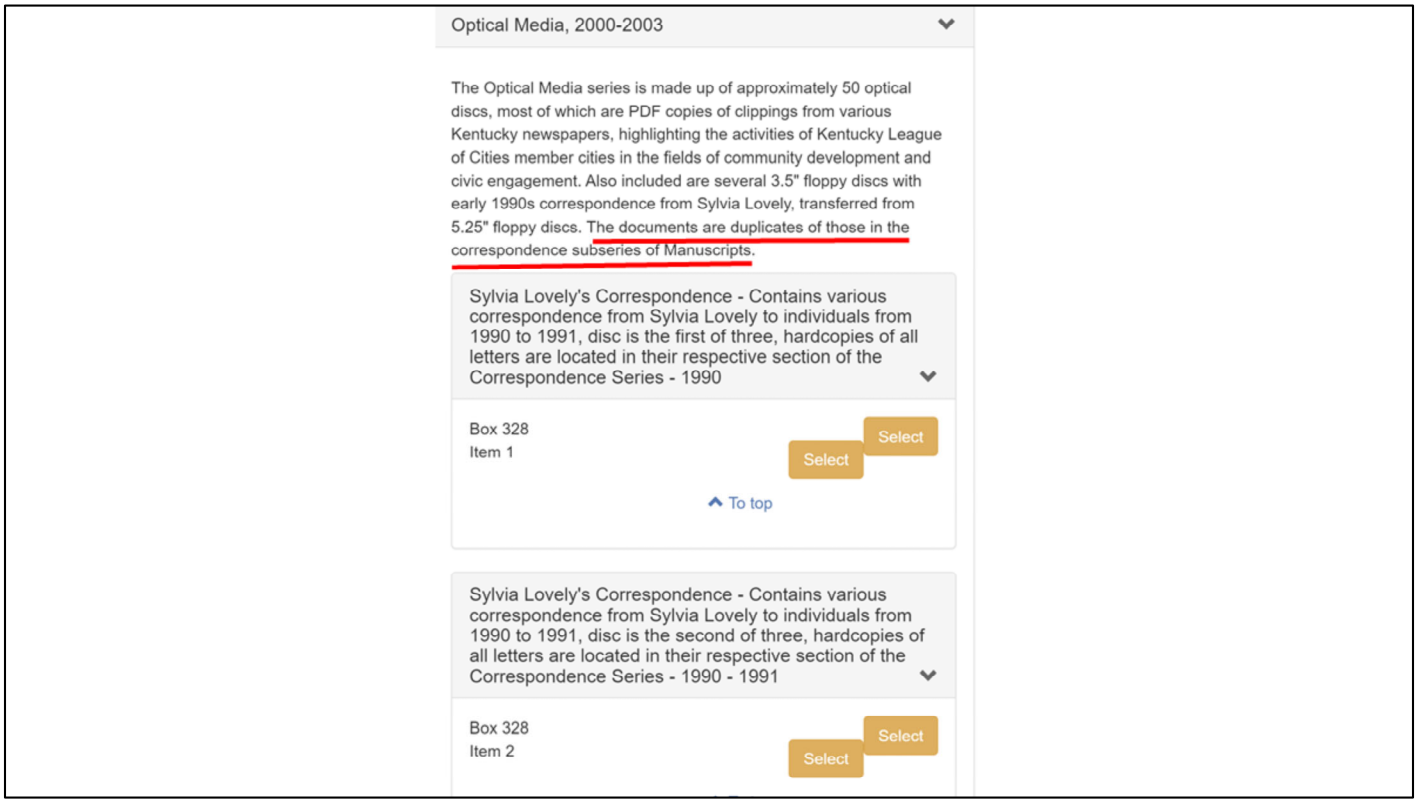
As you can see the description starts out promisingly with a series for Optical Media highlighted in the abstract.

Case Study 1: Kentucky League of Cities records

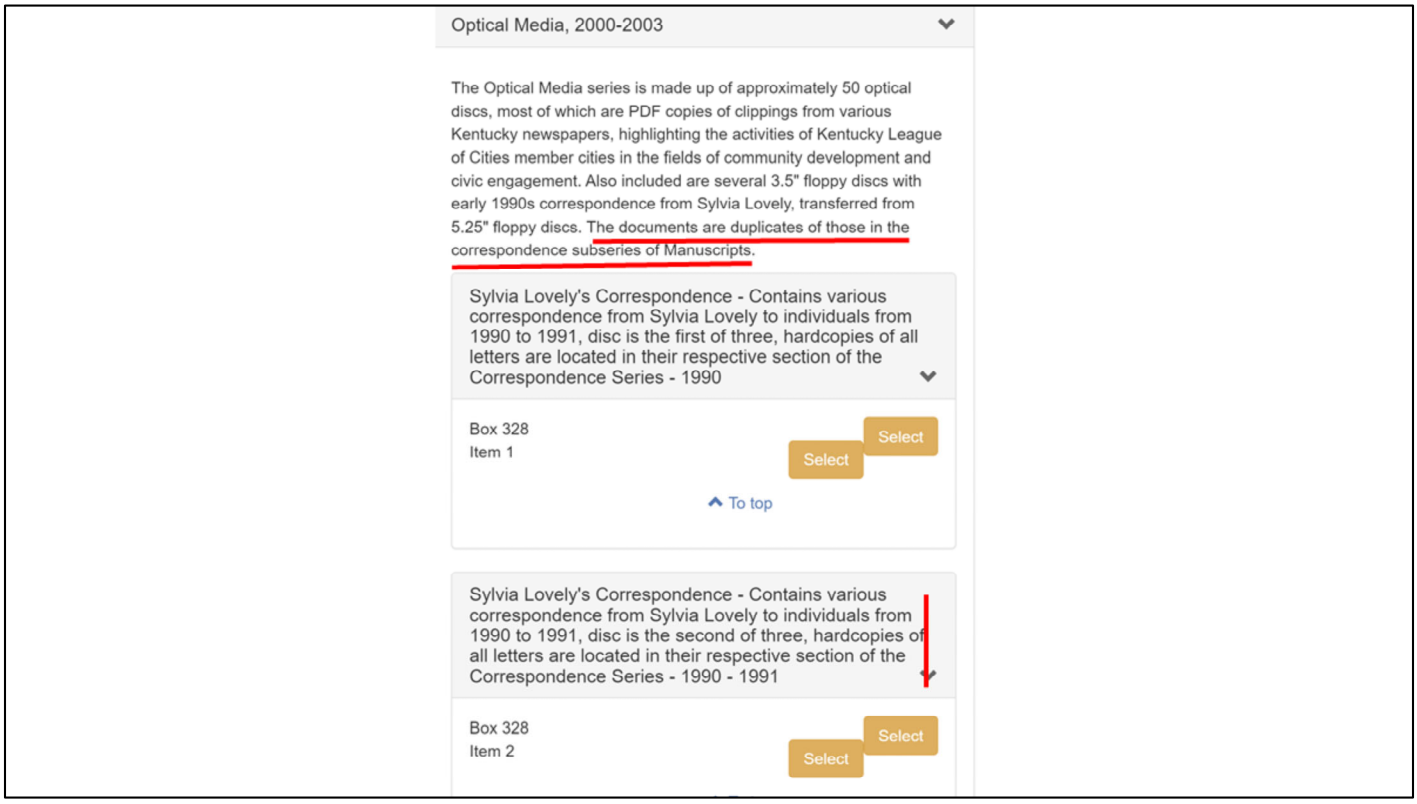
But as you can also see – the extent does not include a byte size or number of digital files.



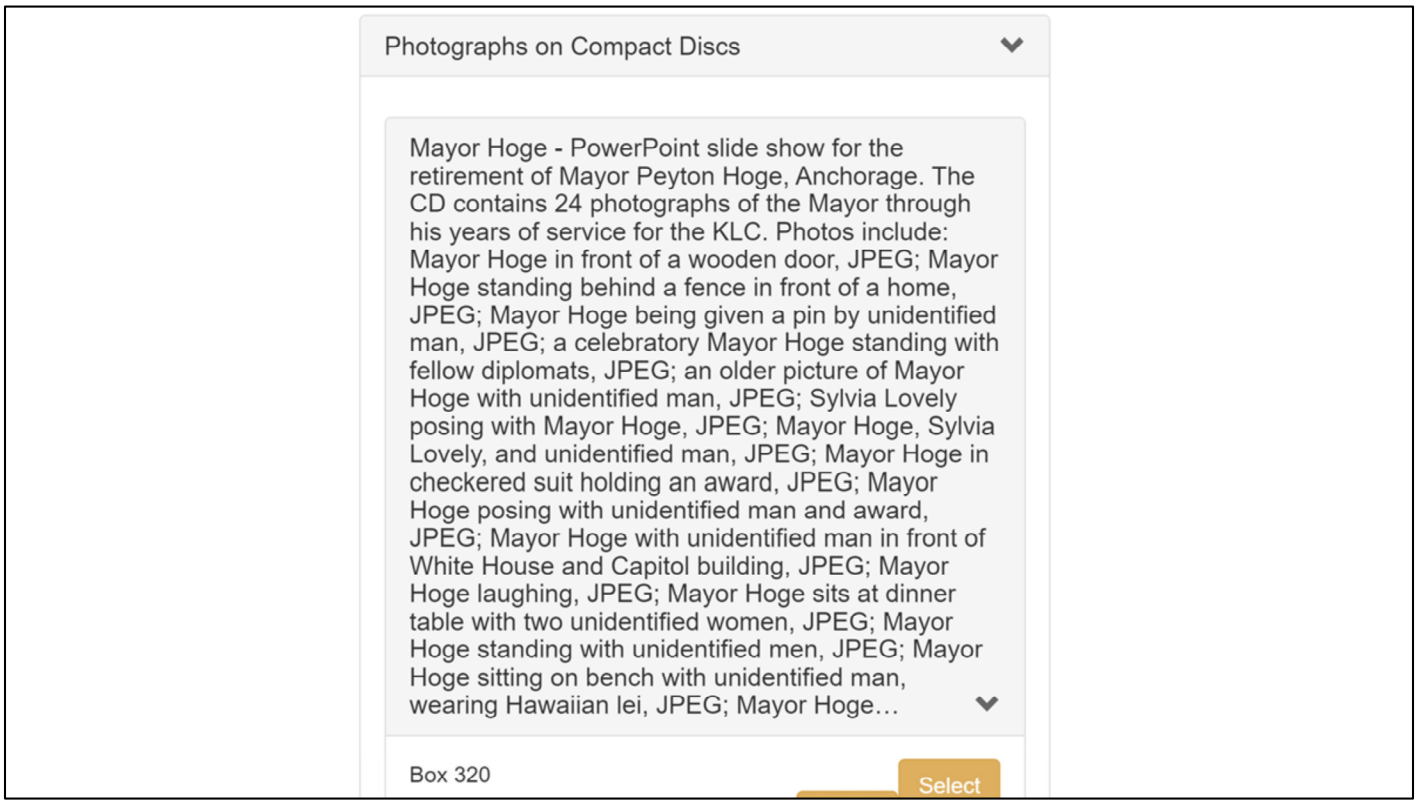
This is a bare bones scope and content note for the collection – but do not worry each individual series is described in depth at that level. But what I do want to highlight here is this helpful note that someone (maybe the archivist or maybe the donor) took the time to actually migrate the files from the 5.25 inch floppies to 3.5 inch floppies! Very forward thinking!



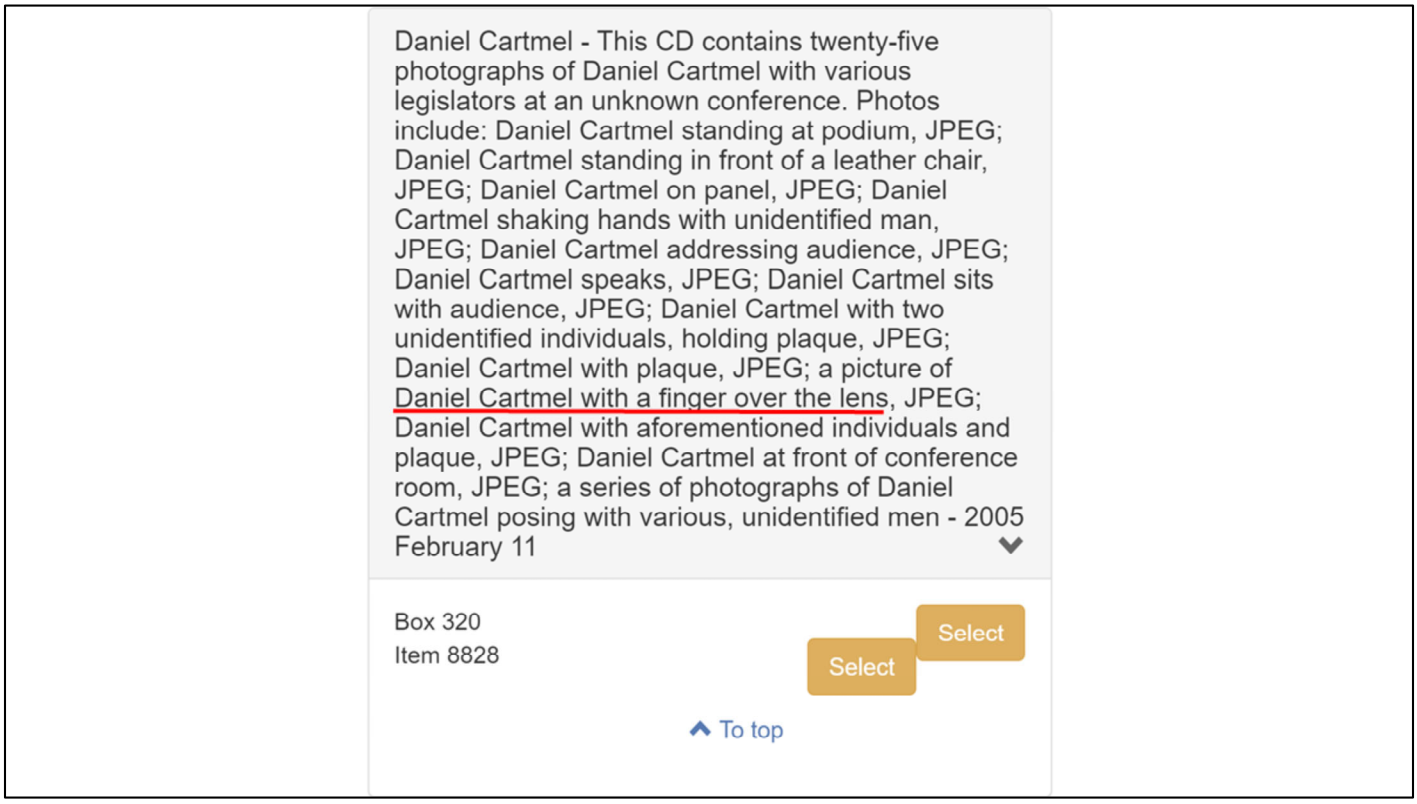
Now the Optical media has their own series and own in depth description – except they have lumped in floppies here (and they are not optical media) and it is not explicitly called out in the file level description which are floppies and which are CDs. I want to note that in 2005 all of these documents were printed out and included in the rest of the finding aid! I do not know if it was by the archivist or the donor, but this does begin to make me question – well then what is the point of keeping the digital files then? Maybe some appraisal should take place. Are these high enough use items with an interesting back history to warrant keeping the digital files? Since these have never been used according to our records – I am guessing probably not.



One other issue in the file list is a muddling between the folder titles and what should ideally be in a scope and contents note. They were really trying to be super helpful and thorough though!



Another group of digital files live as a subseries in the Photographs series. Photographs on Compact Discs! So, as you can see they have lumped a description of every photograph in a PowerPoint slide deck into the title field…This is a fun example because there was so much text in the title it hit our finding aid viewer character limit. Thus, the ellipses. It cuts off – there are over 400 characters missing from this description!



The other discs in this sub series are described in a similar way – if it contained photographs all the photographs are individually described in the title field. This is my favorite example because all of these photographs are of Daniel Cartmel (who that is I do not know) but they are all individually described along with one noting where someone has their finger over the lens!

Case Study 1: KLC Records Issues

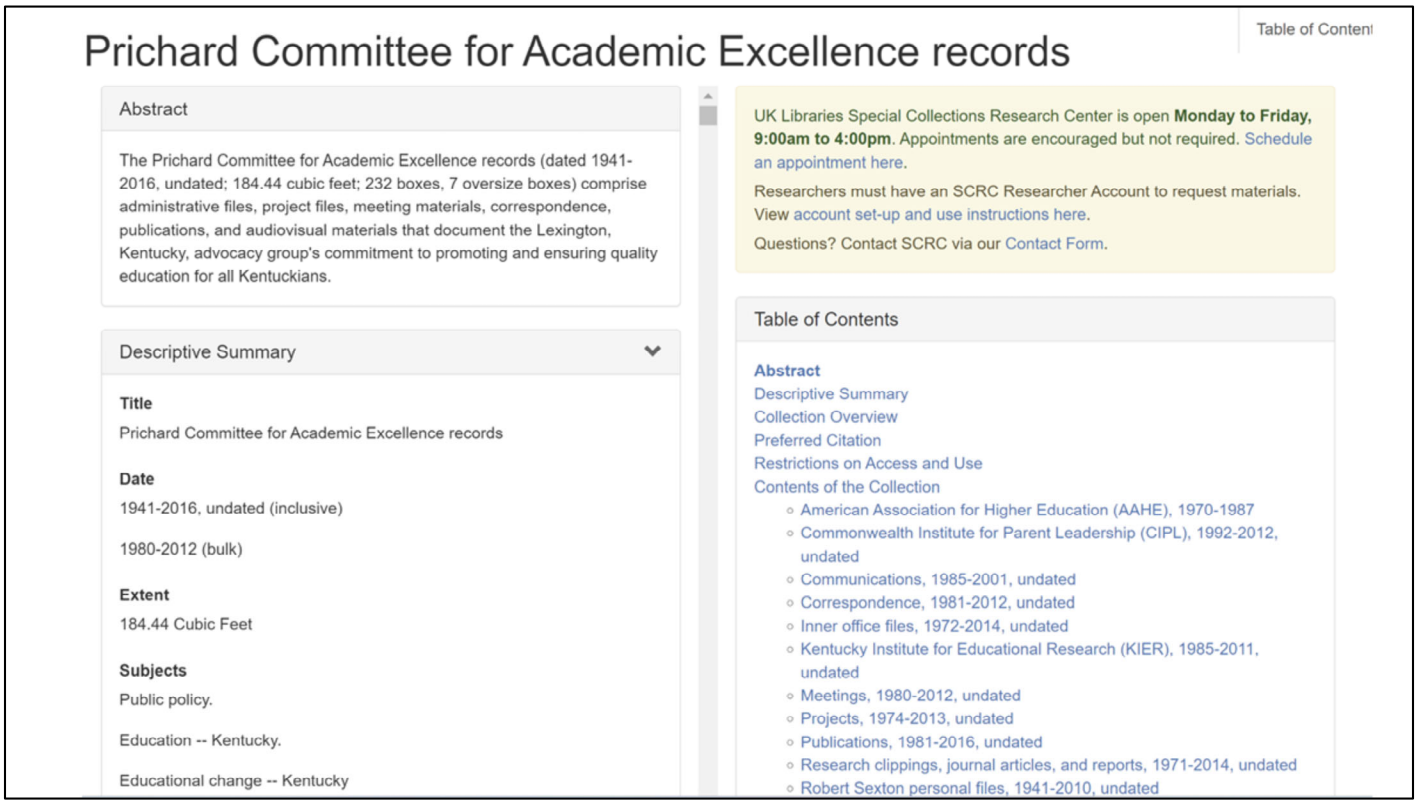
- No extent (bytes or files)
- Description at “CD” level is too granular
- Scope note information in the file titles
- Each individual file or photograph is described in a lump
- Many files are duplicate of existing hard copies
- Individual disk images exist
- Many of the CDs failed in transfer (description still exists for it online)

So, there are a lot of issues I see with this finding aid. There is no extent including the digital footprint (bytes and numbers of files); I think the description at the CD level is too granular for this collection. Scope note information is lumped in with the file titles. Then there is the issues of the files and/or photographs described in a giant mass of text in the title. Many of the text-based files are duplicates of hard copies. We have individual disk images of each CD/floppy which is overkill – and takes up a lot of space (think of all of the space we are paying for and its environmental impact!). Some of the CDs also failed to transfer and the description still exists online for files that are no longer readable. So how do we fix this huge (well-intentioned) mess.

Case Study 1: Roadmap development of born digital description at UKL

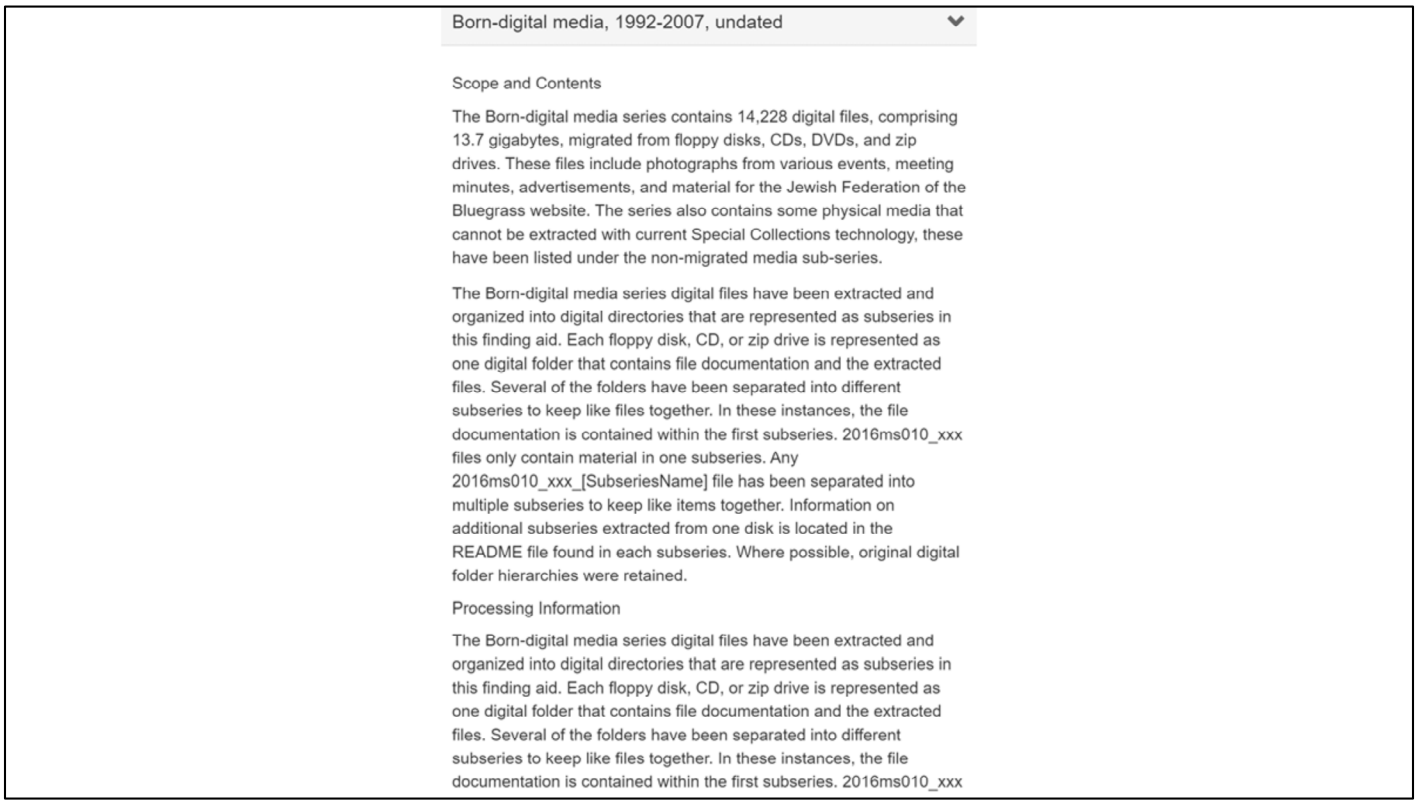
- Prichard Committee records
  - started in 2016, paused in 2017, and completing in 2024
- Jewish Federation of the Bluegrass
  - started and completed in 2017
- Jim Gray mayoral papers
  - started in 2020 and completed in 2021
- Kentucky Quilt Project records
  - started in 2023 and completed in early 2024

So, to get to how we are going to fix this finding aid up…I want to take you on a journey through the UKL's history with the description of born digital records. These are all projects that were formative in our eventual approach to born digital description. Trust me, it will make sense when we get to the end of it. I will talk about the different kinds of things we tried, mistakes we made, and what worked! And end with our documentation and tools for how we describe born digital records today! And you will also be treated to three lovely videos featuring Andrew McDonnell.

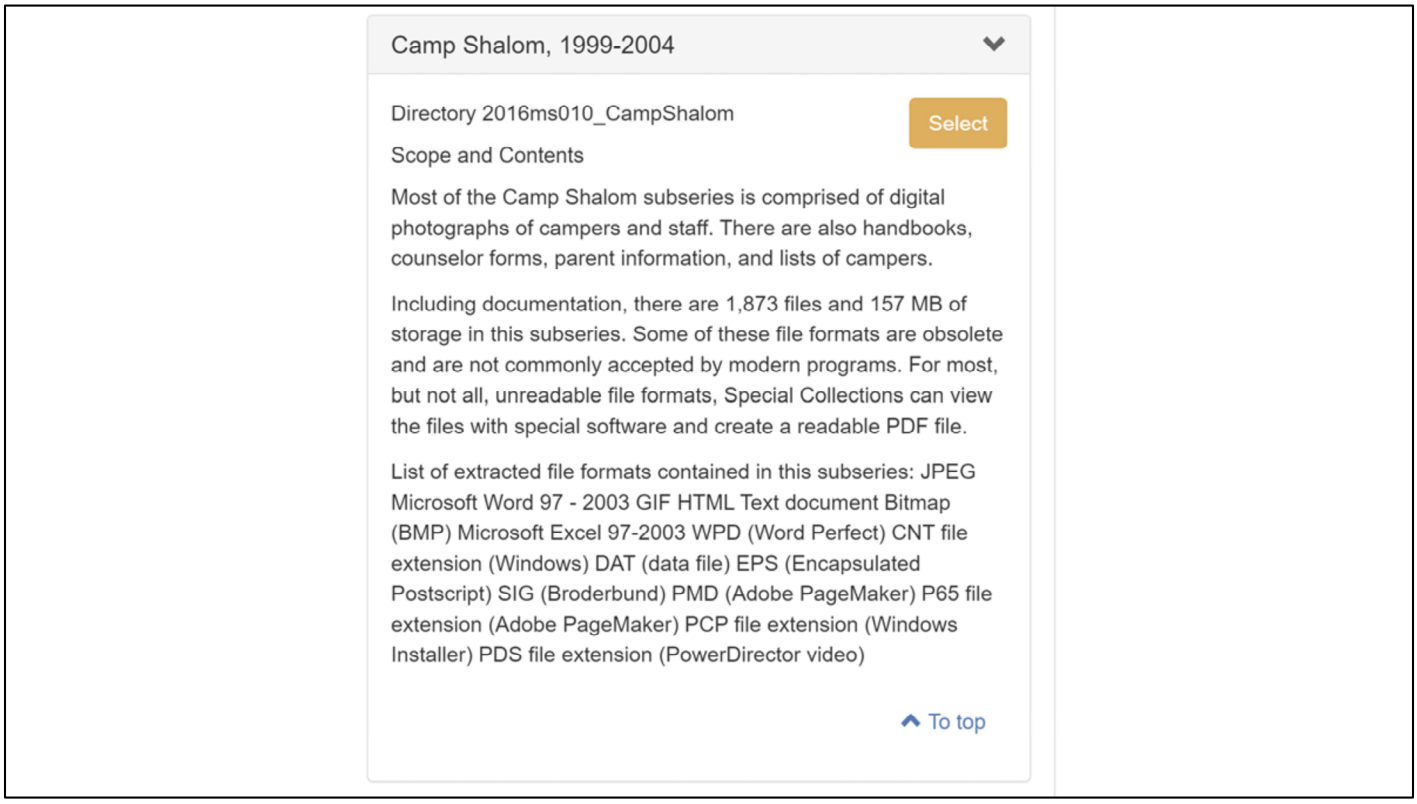


So, this journey will be a little bookended by the Prichard Committee for Academic Excellence records. This was the first collection we used to test out our migration workflows after they were developed around 2015/2016. And we did it intensively, making disk images of everything. And then we got stuck on the description part – absolutely overwhelmed. And the project was paused when the archivist who did the migration work left. We are currently working on this project right now and plan to finish it up this year – but Andrew has a lovely video of this portion of the project later in the presentation. I wanted to keep this one first to show you all – that the description of a massive number of files was our sticking point – and we utterly failed at it initially.

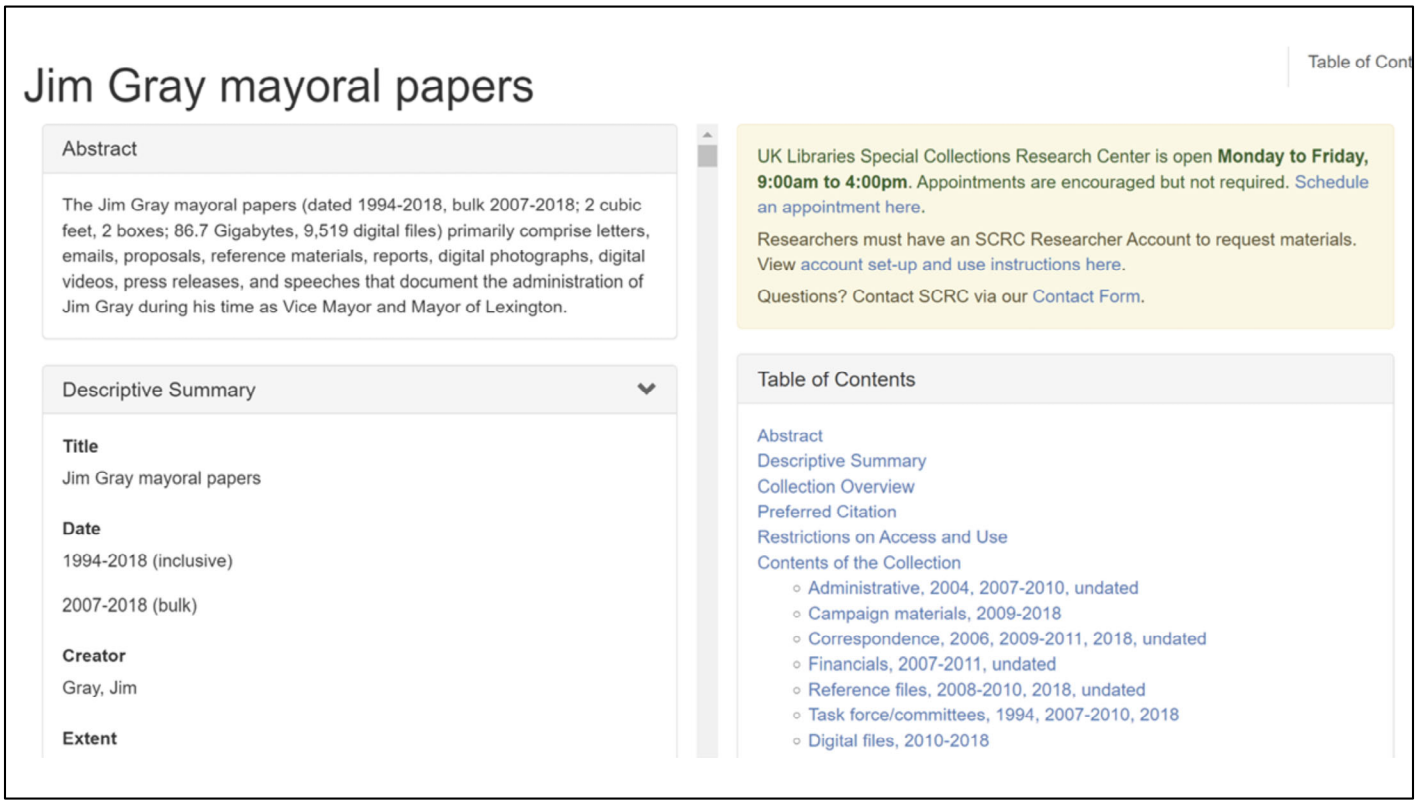
So later in 2017 – we approached the processing of the born digital records attached to the Jewish Federation of the Bluegrass records a little differently. We had a definite timeline with this project as it was grant funded. For this one we backed off from item level and even folder level description. We went for series and sub-series level. All the floppies, CDs, and zip drives were separated into their own Series called born-digital media.



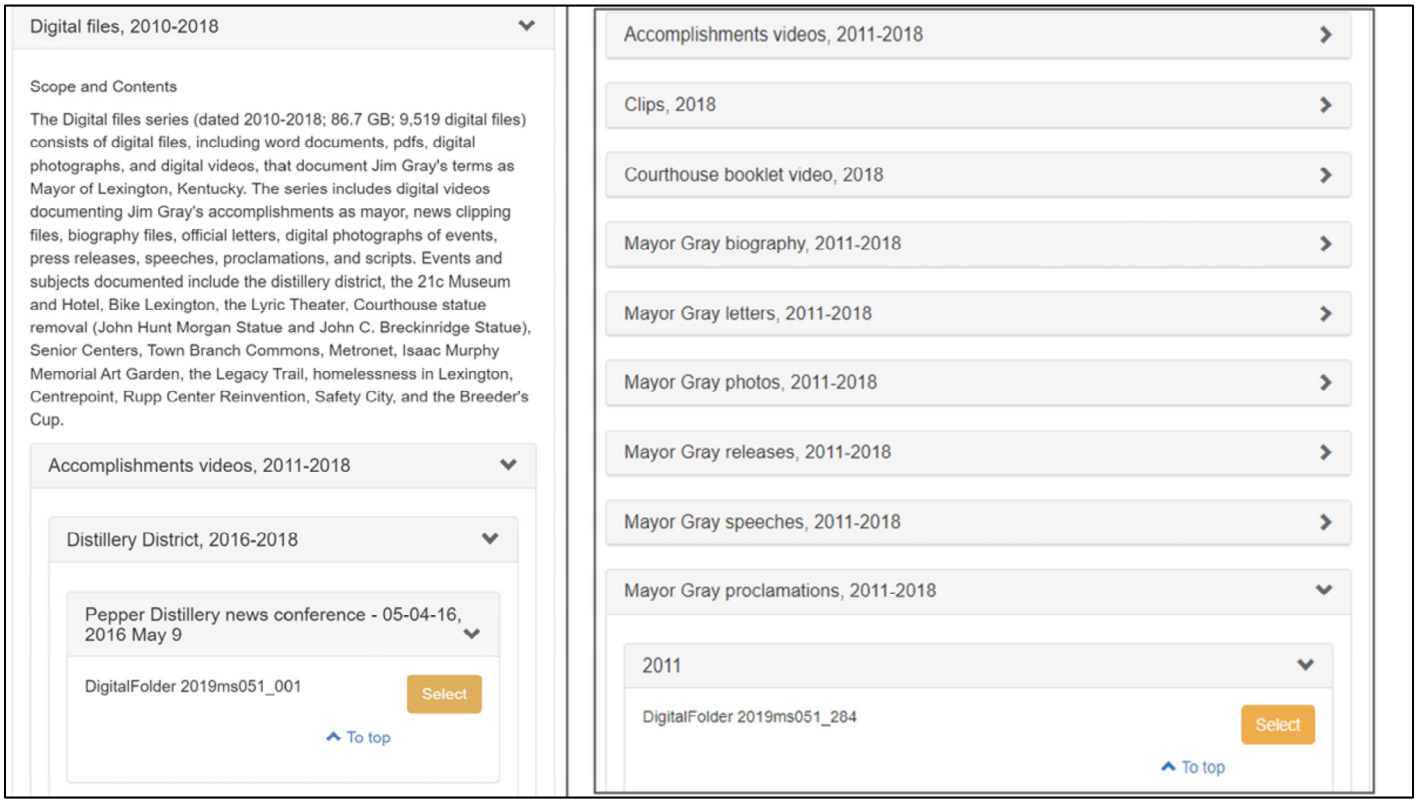
These were then described with an extremely in-depth scope and contents note and processing note about the files, their organization, and what was done to the files upon processing.



Here is one of the sub-series that has all the files associated with Camp Shalom and a very in-depth scope and content note. We went no lower than that in the description and really leaned into the less is more approach with the aim of making the records available quickly.



But then the pandemic happened…and what did we have – so much time on our hands, almost a year's worth of it at home. We used this time to work with some collections that had significant born digital components as they could be completed remotely. The Jim Gray Mayoral papers was one of them. We again went the route of a separate series for the files but provided a listing to the folder level or rather the lowest directory level.



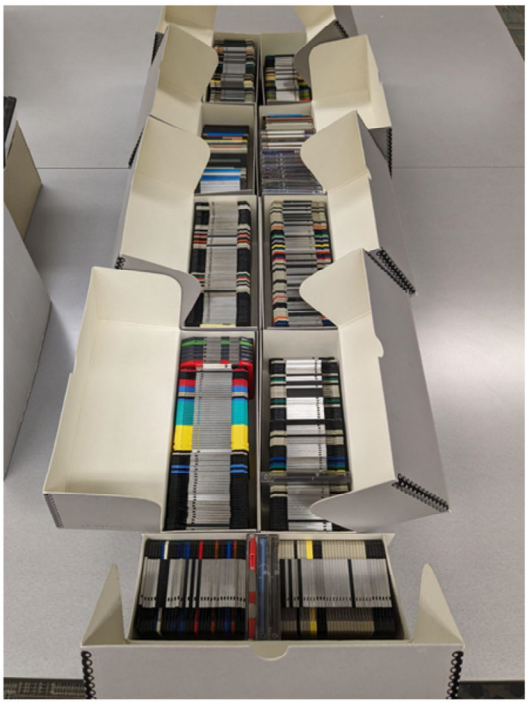
Here you can see a few screenshots of what we did. We provided item level description for the audiovisual materials, but only went to the lowest directory level for the other sub-series. For this collection, there are a lot of local history events in it – so we felt it might be useful to patrons who are looking for information on say when the Pepper Distillery District opened in Lexington or the folder of photographs for the 2011 Christmas tree lighting. We made this decision to go lower partially because of the pandemic but also because the files were impeccably arranged, and they had excellent, human readable file and directory names. But it was a lot of work – and because of global events we had the time. We will not always have that though.

Prichard Reappraisal by the Numbers

Prichard Pre-Reappraisal: 33.1 GB
4,668 Directories
43,810 Files

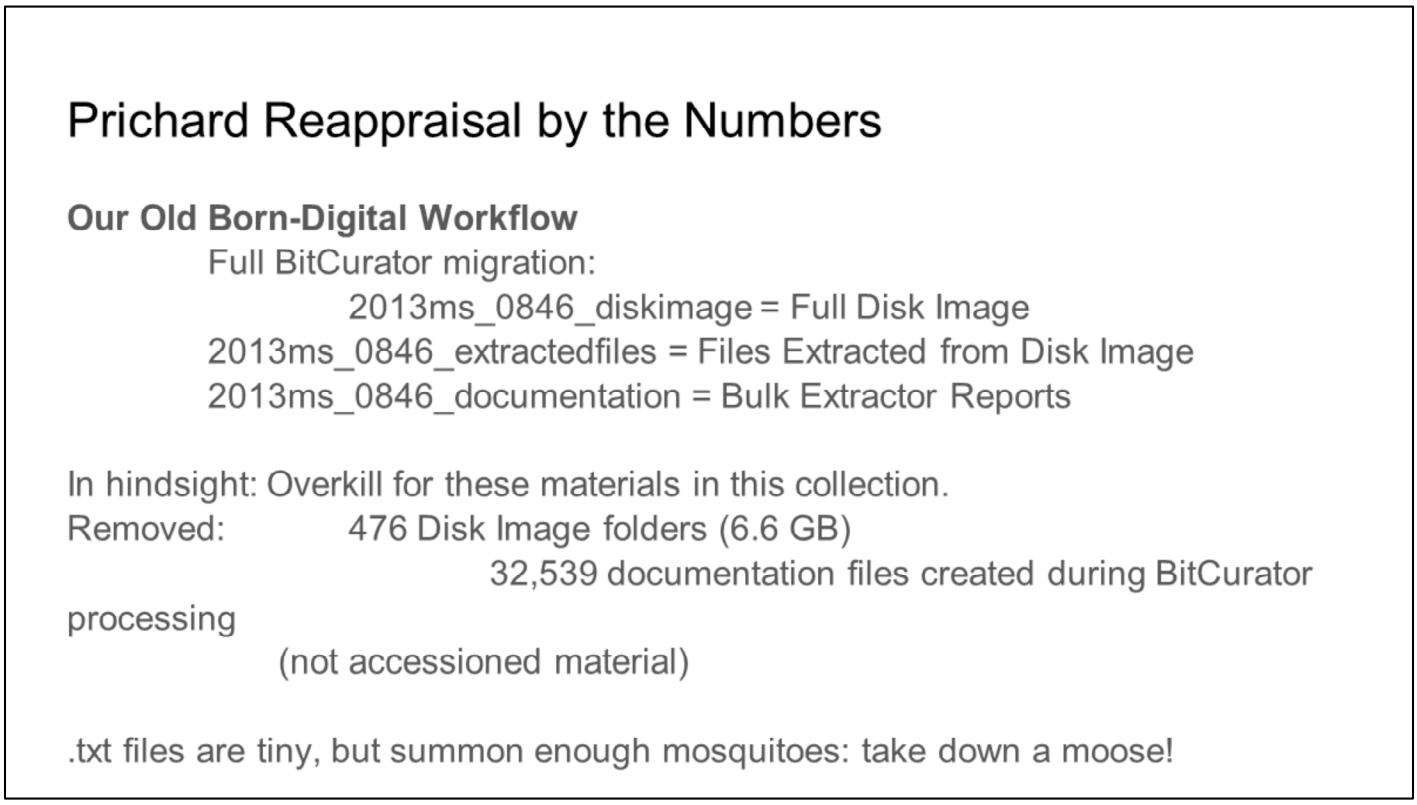
Prichard After Reappraisal: 22 GB
1,206 Directories
9,570 Files

Removed ⅓ of our Data. Cool.
But ⅘ of our files? How and why?



The Prichard Committee for Academic Excellence is proving to be a valuable collection in part because its physical material is so thoroughly described. As a result it gives us one model to help establish how we want to handle description of a hybrid collection of this size and character. It's also an excellent illustration of how we processed born-digital material from floppy disks at the start of this effort, and how we might re-work that methodology.

What I found already migrated to our networked servers when this was first processed 8 years ago, was 33.1 GB of data. 43,810 Files in 4,668 Directories. 33 Gigabytes isn't a daunting quantity to store, but 4,668 folders is absolutely a daunting quantity to describe. After digging into what was actually in the migrated data, a very quick intervention allowed us to remove over 3,400 directories and 34,000 files. 80% of the files, gone, with a clean conscience. Why? Because we were the ones to create those extra tens of thousands of files.

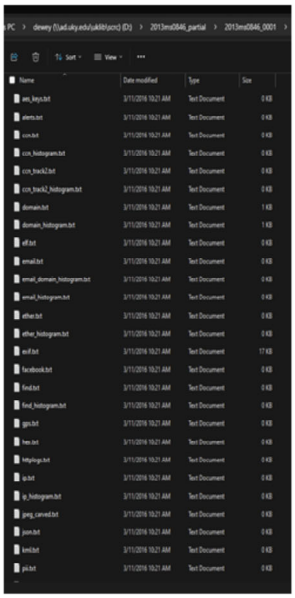


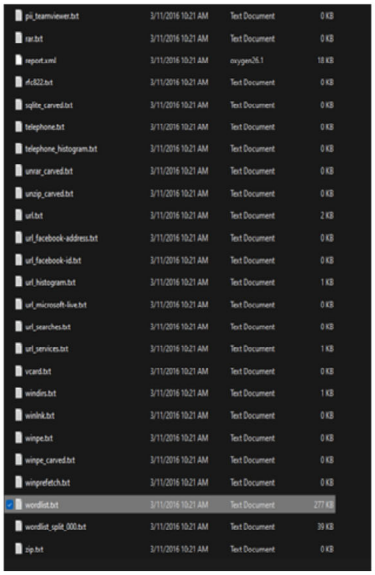
The archivist that processed these materials had done an incredible job, so don't mistake this for criticism. But in implementing a full forensic migration, they had created more metadata and supplemental documentation than we decided the collection warranted.

For each individual floppy disk, they virus scanned the disk, conducted a full transfer by creating a disk image for the floppy. They analyzed the contents of the disk image using Bulk Extractor to search for PII, and then finally, extracted all the files from the disk images.

On re-appraisal, we determined the disk images were not worth the digital space they occupied on our servers. If this had been a higher value collection, the duplication of the disk image and the files extracted would have been worth preserving. But we decided to just keep the extracted files and delete the disk images.

Prichard Reappraisal by the Numbers





.txt files are tiny, but enough mosquitoes can: take down a moose! (As we say in Kentucky)

Likewise, when we looked in the Documentation folders for each disk, we found text logs for each search Bulk Extractor ran on each individual disk. For instance, it did a scan for email addresses or credit card numbers and created a txt file for each search, even if it found nothing. However, it ran those searches for almost 50 categories for each of 694 disks. The end result was over 34,000 tiny txt files that could be removed. Most of them contained absolutely no information, and a PII scan encompassing the full collection could be more efficiently conducted after our reappraisal was concluded.

Prichard Re-Description

Previous:

Only unmigrated materials mentioned.

Migrated materials drifting in the description vacuum of space.

Now:

Directory-level, integrated description

CIPL lists, 1998	File	mixed materials	Box: 2013ms0846-013, Folder: 15
Correspondence, 1998	File	mixed materials, mixed materials, ...	Box: 2013ms0846-011, Folder: 24; Box
Curriculum outline, 1998	File	mixed materials	Box: 2013ms0846-011, Folder: 25
Danforth Foundation report, 1998	File	Digital File	DigitalFolder: 2013ms0846_borndigital,
Design forms, 1998	File	mixed materials	Box: 2013ms0846-011, Folder: 26
Elizabethtown, 1998	File	mixed materials	Box: 2013ms0846-012, Folder: 4
Georgetown, 1998	File	mixed materials	Box: 2013ms0846-012, Folder: 5
Georgetown portraits, 1998	File	Digital File	DigitalFolder: 2013ms0846_borndigital,

So, after this clean-up, while the description task was not small, it was 80% smaller than it had been.

Previously, the Prichard collection guide only mentioned unmigrated materials. Migrated materials were left drifting in the description vacuum of space. After reappraisal, our description will be folder/or directory-level description, integrated into the description of the physical objects in the collection.

Prichard Re-Description

Why?

In this case, we already had a thoroughly processed collection with over 5,100 physical folders spread across 232 boxes. The intellectual arrangement was already there and sped the process of describing the born-digital material. Pre-existing Series and Sub-series could accommodate all the born-digital materials, in this case described at the directory/folder level.

CIPL lists, 1998	File	mixed materials	Box: 2013ms0846-013, Folder: 15
Correspondence, 1998	File	mixed materials, mixed materials, ...	Box: 2013ms0846-011, Folder: 24; Box
Curriculum outline, 1998	File	mixed materials	Box: 2013ms0846-011, Folder: 25
Danforth Foundation report, 1998	File	Digital File	DigitalFolder: 2013ms0846_borndigital,
Design forms, 1998	File	mixed materials	Box: 2013ms0846-011, Folder: 26
Elizabethtown, 1998	File	mixed materials	Box: 2013ms0846-012, Folder: 4
Georgetown, 1998	File	mixed materials	Box: 2013ms0846-012, Folder: 5
Georgetown portraits, 1998	File	Digital File	DigitalFolder: 2013ms0846_borndigital,

Why? Well, in this case, we already had a thoroughly processed collection with over 5,100 physical folders spread across 232 boxes. The intellectual arrangement very much reflected the contents we found in the digital records. So not only was it appropriate to keep the physical and digital intellectually united, but the enormous series and sub-series framework sped the process of describing the born-digital material.

I won't always be so lucky, when it comes to walking into born-digital description where the outline of the puzzle is already finished. But in this instance it made the most sense and helped define a criteria for how and when to describe born-digital materials in this fashion.

Kentucky Quilt Project: Now with File Manifests

So unassuming.

As noted last year: we started with 146,664 files; 3.86 TB

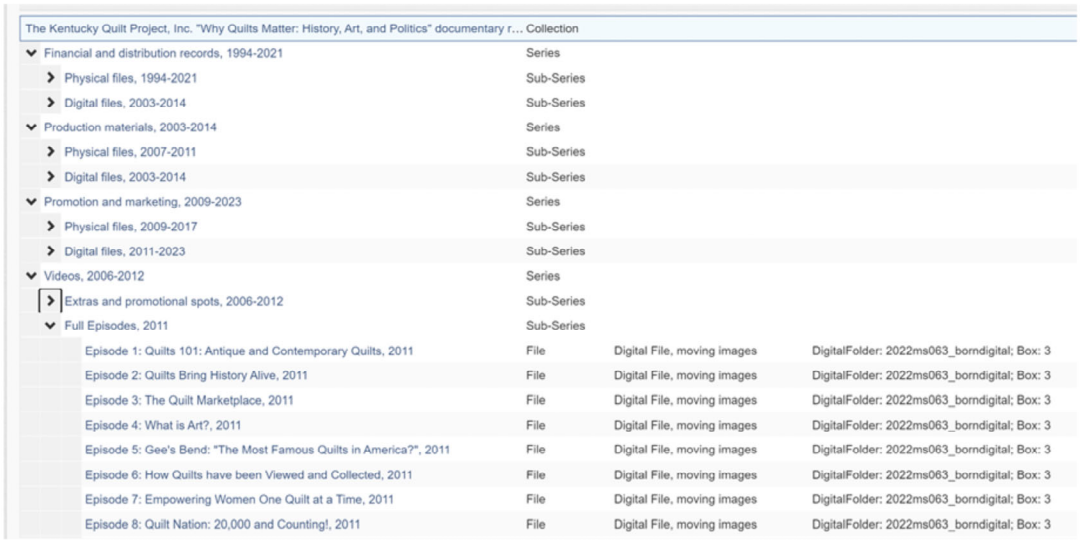
We conclude with: 3,006 files; 418.6 GB

But how best to describe 3,006 files?

As noted at last year's Best Practices presentation: we started with almost 150,000 files
After deduplication and appraising which of these materials were truly in scope for the collection, We concluded with: 3,006 files
But how best to describe these 3,006 files?

Describe it.

We call it: Tier 2, Semi-Integrated with a File Manifest Twist



The Kentucky Quilt Project, Inc. "Why Quilts Matter: History, Art, and Politics" documentary r…	Collection		
Financial and distribution records, 1994-2021	Series		
Physical files, 1994-2021	Sub-Series		
Digital files, 2003-2014	Sub-Series		
Production materials, 2003-2014	Series		
Physical files, 2007-2011	Sub-Series		
Digital files, 2003-2014	Sub-Series		
Promotion and marketing, 2009-2023	Series		
Physical files, 2009-2017	Sub-Series		
Digital files, 2011-2023	Sub-Series		
Videos, 2006-2012	Series		
Extras and promotional spots, 2006-2012	Sub-Series		
Full Episodes, 2011	Sub-Series		
Episode 1: Quilts 101: Antique and Contemporary Quilts, 2011	File	Digital File, moving images	DigitalFolder: 2022ms063_borndigital; Box: 3
Episode 2: Quilts Bring History Alive, 2011	File	Digital File, moving images	DigitalFolder: 2022ms063_borndigital; Box: 3
Episode 3: The Quilt Marketplace, 2011	File	Digital File, moving images	DigitalFolder: 2022ms063_borndigital; Box: 3
Episode 4: What is Art?, 2011	File	Digital File, moving images	DigitalFolder: 2022ms063_borndigital; Box: 3
Episode 5: Gee's Bend: "The Most Famous Quilts in America?", 2011	File	Digital File, moving images	DigitalFolder: 2022ms063_borndigital; Box: 3
Episode 6: How Quilts have been Viewed and Collected, 2011	File	Digital File, moving images	DigitalFolder: 2022ms063_borndigital; Box: 3
Episode 7: Empowering Women One Quilt at a Time, 2011	File	Digital File, moving images	DigitalFolder: 2022ms063_borndigital; Box: 3
Episode 8: Quilt Nation: 20,000 and Counting!, 2011	File	Digital File, moving images	DigitalFolder: 2022ms063_borndigital; Box: 3

We call it: Tier 2, Semi-Integrated with a File Manifest Twist

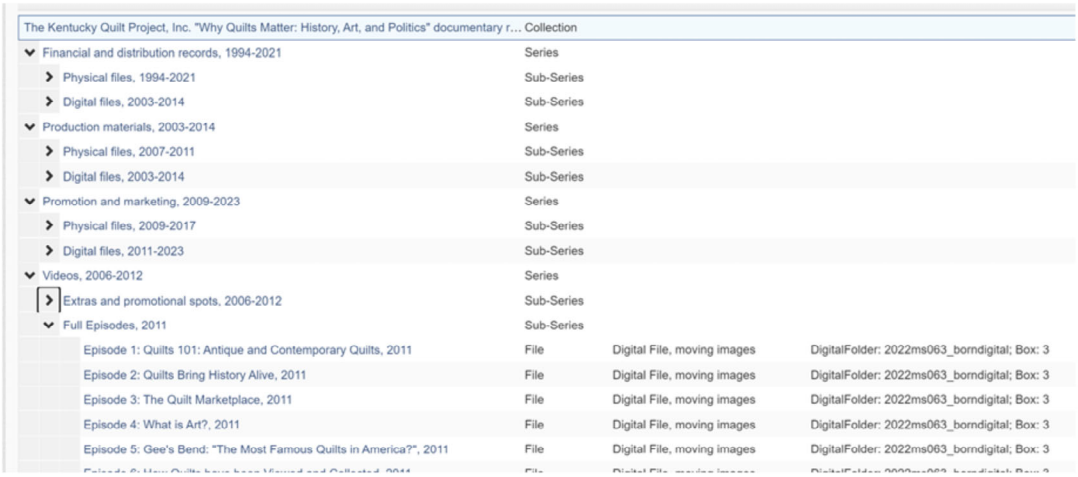
Well, we don’t actually call it that, but that’s essentially the solution we settled on.Practically speaking, we have a hybrid collection, fairly evenly split between physical and digital materials, strictly based on the number of individual items in each category.

In addition, the physical and digital materials fall quite nicely into the same intellectual arrangement. As you can see here we have four primary series (Financial, Production, Promotion, and the actual heart of the collection, Videos).

Describe it.

Beyond digital Sub-series:

Item level description for largest, essential records in the collection: final video files for full documentary episodes



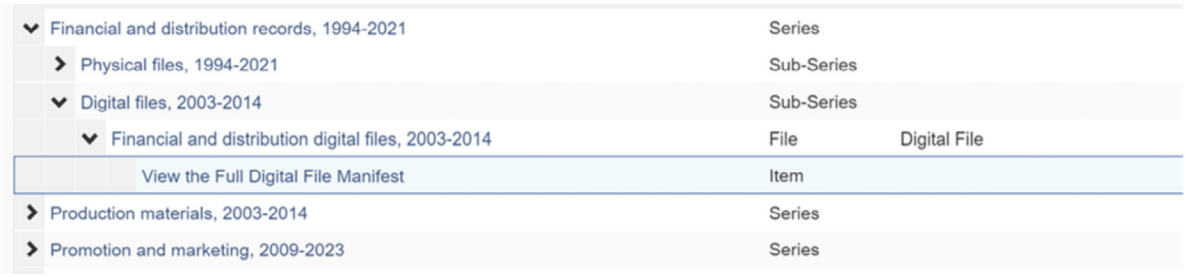
The Kentucky Quilt Project, Inc. "Why Quilts Matter: History, Art, and Politics" documentary r…	Collection		
Financial and distribution records, 1994-2021	Series		
Physical files, 1994-2021	Sub-Series		
Digital files, 2003-2014	Sub-Series		
Production materials, 2003-2014	Series		
Physical files, 2007-2011	Sub-Series		
Digital files, 2003-2014	Sub-Series		
Promotion and marketing, 2009-2023	Series		
Physical files, 2009-2017	Sub-Series		
Digital files, 2011-2023	Sub-Series		
Videos, 2006-2012	Series		
Extras and promotional spots, 2006-2012	Sub-Series		
Full Episodes, 2011	Sub-Series		
Episode 1: Quilts 101: Antique and Contemporary Quilts, 2011	File	Digital File, moving images	DigitalFolder: 2022ms063_borndigital; Box: 3
Episode 2: Quilts Bring History Alive, 2011	File	Digital File, moving images	DigitalFolder: 2022ms063_borndigital; Box: 3
Episode 3: The Quilt Marketplace, 2011	File	Digital File, moving images	DigitalFolder: 2022ms063_borndigital; Box: 3
Episode 4: What is Art?, 2011	File	Digital File, moving images	DigitalFolder: 2022ms063_borndigital; Box: 3
Episode 5: Gee's Bend: "The Most Famous Quilts in America?", 2011	File	Digital File, moving images	DigitalFolder: 2022ms063_borndigital; Box: 3

Given the anticipated research value of most of those materials, which was relatively low, we decided to just provide a sub-series level description of most of the digital materials and folder level description of the physical. The exception to this was item level description for the finalized video episodes of the documentary, the focus of this collection.

Taking the time to manually author description for the rest of the digital materials was deemed beyond the requirements of this collection, given the time it would take relative to its value. But it was an excellent candidate for automated description with a simple, interactive file manifest.

Describe it.

Visitors are able to open a navigable File Manifest in a new tab.

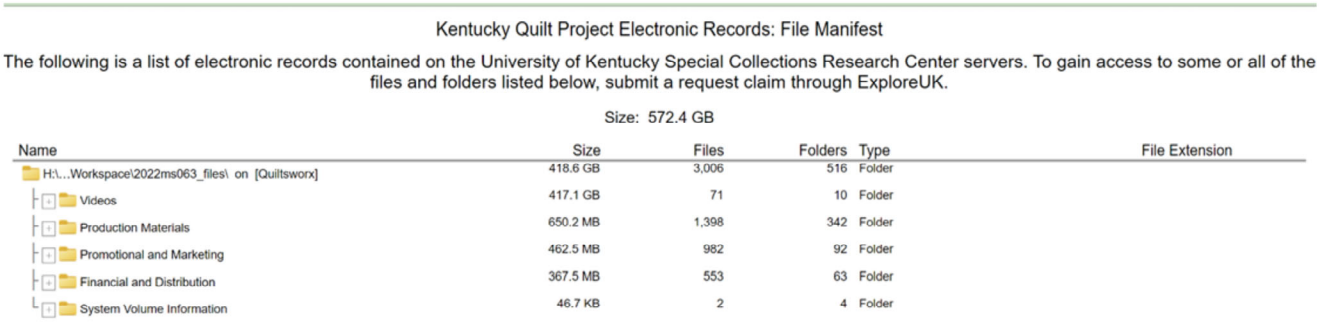


We liked the idea, presented in description samples by institutions such as the University of Minnesota and recommended by the University of California, of offering a full file manifest for exactly this type of collection. The Kentucky Quilt Project had a large number of files to describe and its folder and file titles were very human readable, they would genuinely mean something to a researcher. And generating a file manifest would take nearly no time at all while still providing valuable description to a researcher.

TreeSize generated HTML File Manifest

The Kentucky Quilt Project lent itself well to file manifest viewing:

- Large number of files
- Human readable file and folder titles
- Cost/Benefit analysis = Worth generating folder/file manifest but not worth time necessary to describe digital materials in resource record at this granularity



Kentucky Quilt Project Electronic Records: File Manifest

The following is a list of electronic records contained on the University of Kentucky Special Collections Research Center servers. To gain access to some or all of the files and folders listed below, submit a request claim through ExploreUK.

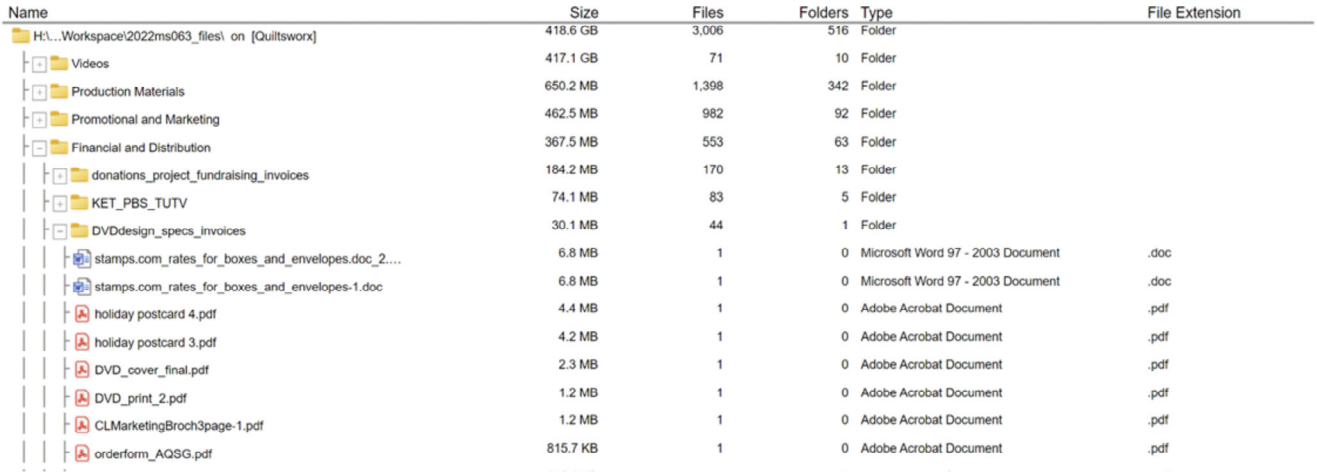
Size: 572.4 GB

Name	Size	Files	Folders	Type	File Extension
H:\...Workspace\2022ms063_files\ on [Quiltsworx]	418.6 GB	3,006	516	Folder	
Videos	417.1 GB	71	10	Folder	
Production Materials	650.2 MB	1,398	342	Folder	
Promotional and Marketing	462.5 MB	982	92	Folder	
Financial and Distribution	367.5 MB	553	63	Folder	
System Volume Information	46.7 KB	2	4	Folder	

Rather than provide a 3,000 row spreadsheet, we thought it might be interesting to use the HTML file manifest we are able to generate using TreeSize, a tool we already use for appraising, describing, deduplicating, and generally understanding our digital materials.

TreeSize generated HTML File Manifest

Collapsible and expandable. But will it confuse researchers expecting full online access?



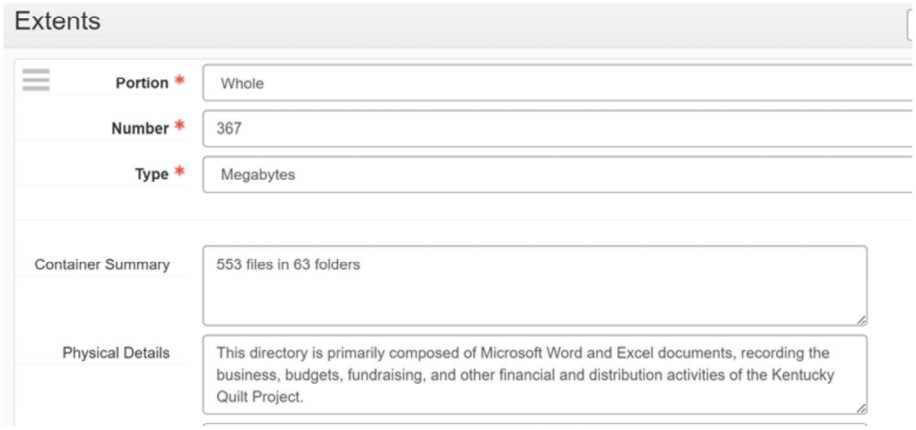
Name	Size	Files	Folders	Type	File Extension
H:\...Workspace\2022ms063_files\ on [Quiltsworx]	418.6 GB	3,006	516	Folder	
Videos	417.1 GB	71	10	Folder	
Production Materials	650.2 MB	1,398	342	Folder	
Promotional and Marketing	462.5 MB	982	92	Folder	
Financial and Distribution	367.5 MB	553	63	Folder	
donations_project_fundraising_invoices	184.2 MB	170	13	Folder	
KET_PBS_TUTV	74.1 MB	83	5	Folder	
DVDdesign_specs_invoices	30.1 MB	44	1	Folder	
stamps.com_rates_for_boxes_and_envelopes.doc_2....	6.8 MB	1	0	Microsoft Word 97 - 2003 Document	.doc
stamps.com_rates_for_boxes_and_envelopes-1.doc	6.8 MB	1	0	Microsoft Word 97 - 2003 Document	.doc
holiday postcard 4.pdf	4.4 MB	1	0	Adobe Acrobat Document	.pdf
holiday postcard 3.pdf	4.2 MB	1	0	Adobe Acrobat Document	.pdf
DVD_cover_final.pdf	2.3 MB	1	0	Adobe Acrobat Document	.pdf
DVD_print_2.pdf	1.2 MB	1	0	Adobe Acrobat Document	.pdf
CLMarketingBroch3page-1.pdf	1.2 MB	1	0	Adobe Acrobat Document	.pdf
orderform_AQSG.pdf	815.7 KB	1	0	Adobe Acrobat Document	.pdf

The HTML manifest provides a collapsible and expandable means of browsing the files and directories as we have them on our servers. Working with Library IT, we received permission and space to host these manifests on a server separate from our finding aids and link out to them from within the relevant finding aid.

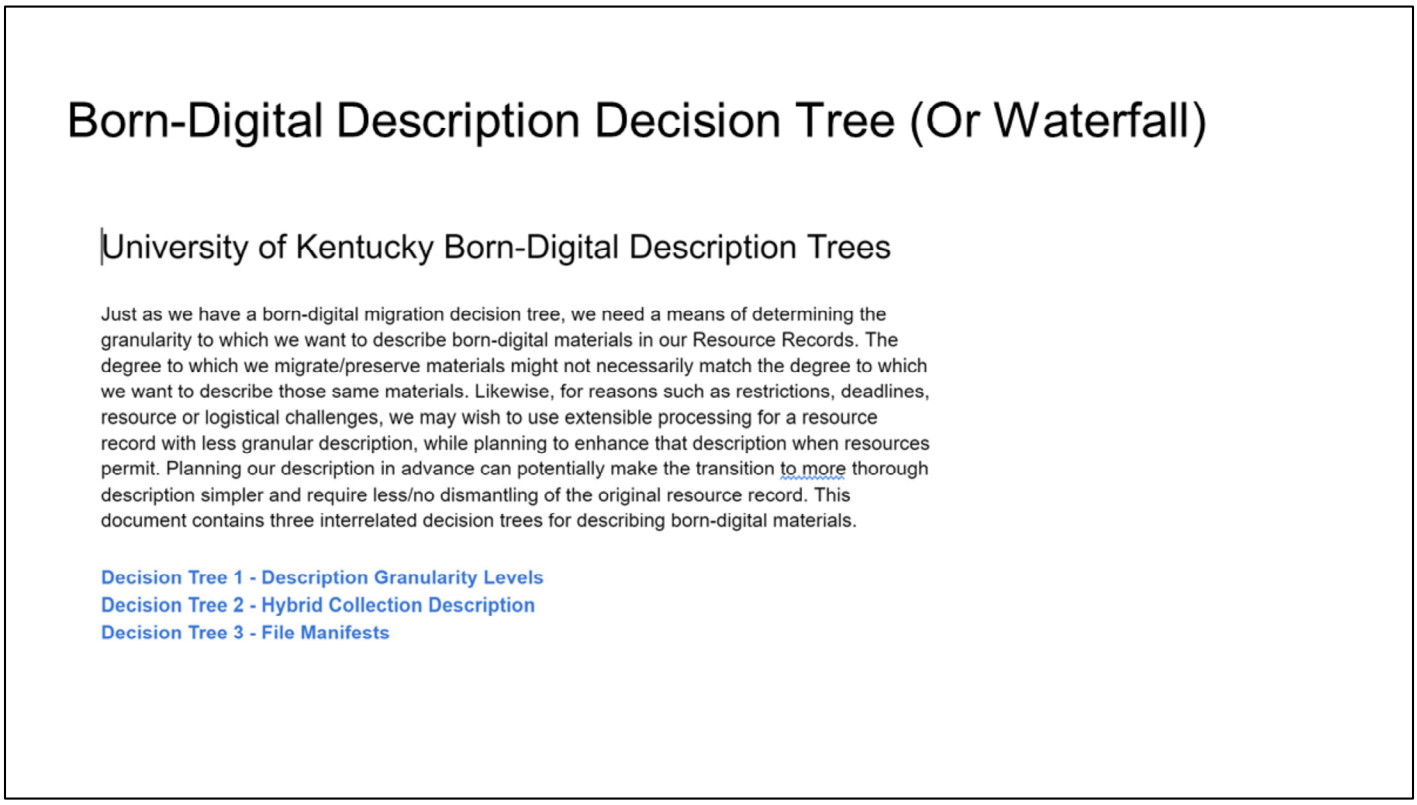
Describe it.

Not *just* File Manifest, though.

Summative extent note reflecting size, quantity, file format, and nature of electronic records in each Sub-series.



In addition, it's worth noting the finding aid also includes a Summative extent note indicating the size, quantity, file format, and nature of electronic records in each Sub-series.

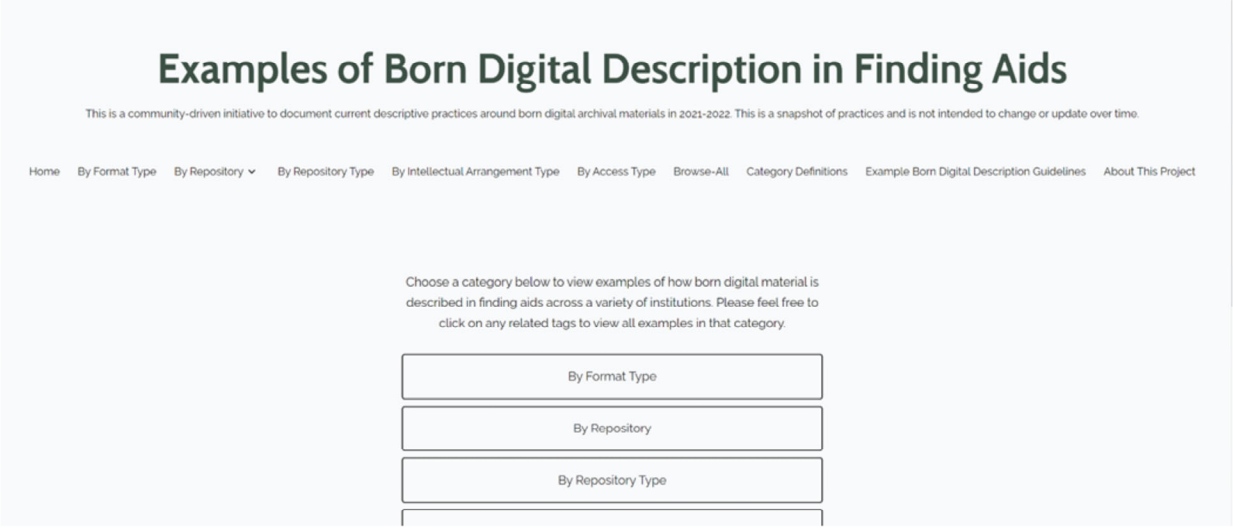


Born-Digital Description Decision Tree

Every time we confronted a new collection or discovered a previously undescribed digital addition to a physical collection (a box of floppy disks, a tower of CDs) the most logical means of describing it deviated from our previous methods in some small way. And existing born-digital material had been described in a number of different ways in previous collection descriptions. So we decided to attempt to codify the ways and granularity to which we would describe different materials to help simplify our approach and make the end result consistent across our Resource Guides.

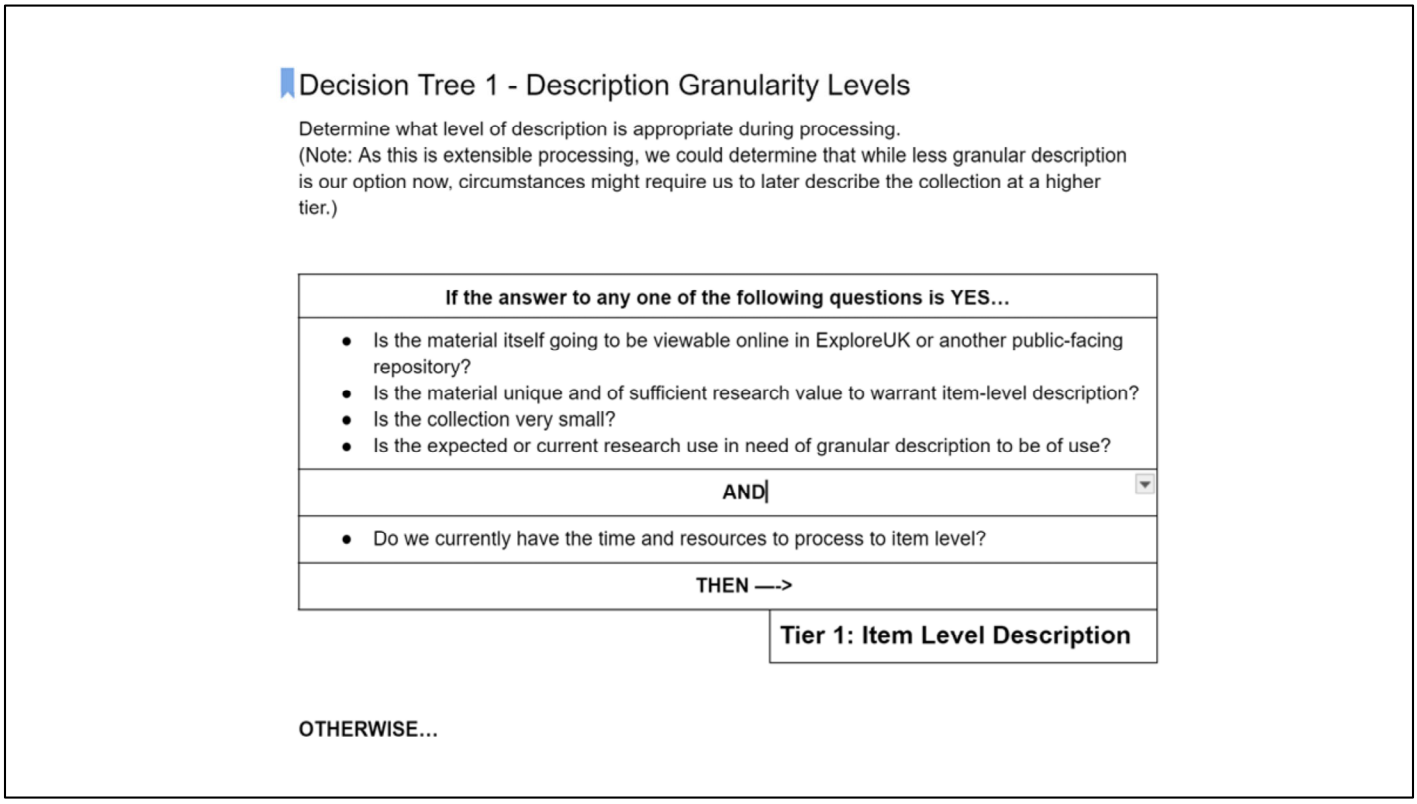
Easier said than done.

Born-Digital Description Decision Tree (Or Waterfall)



We're indebted to the authors of the clearly titled Wordpress site, Examples of Born Digital Description in Finding Aids, and the institutions that contributed their policies and examples, especially the University of Minnesota, The University of California System, and Yale University. Their examples provided an articulate framework to help us more fully define a decision tree for describing our materials.

Ours is truly a living document that will doubtless evolve as we encounter materials that defy our current criteria or require a new level.



One thing that became apparent as we defined classifications was that it wasn't going to be as simple as delineating resource record levels of description into just Item Level, Series Level, Collection Level description across different materials. The first primarily born-digital collection I started with provided an immediate example of the need for something slightly more complex..

In descending order of detailed description, here are our current tiers:

Tier 1 is collections that are fully Item Level Description

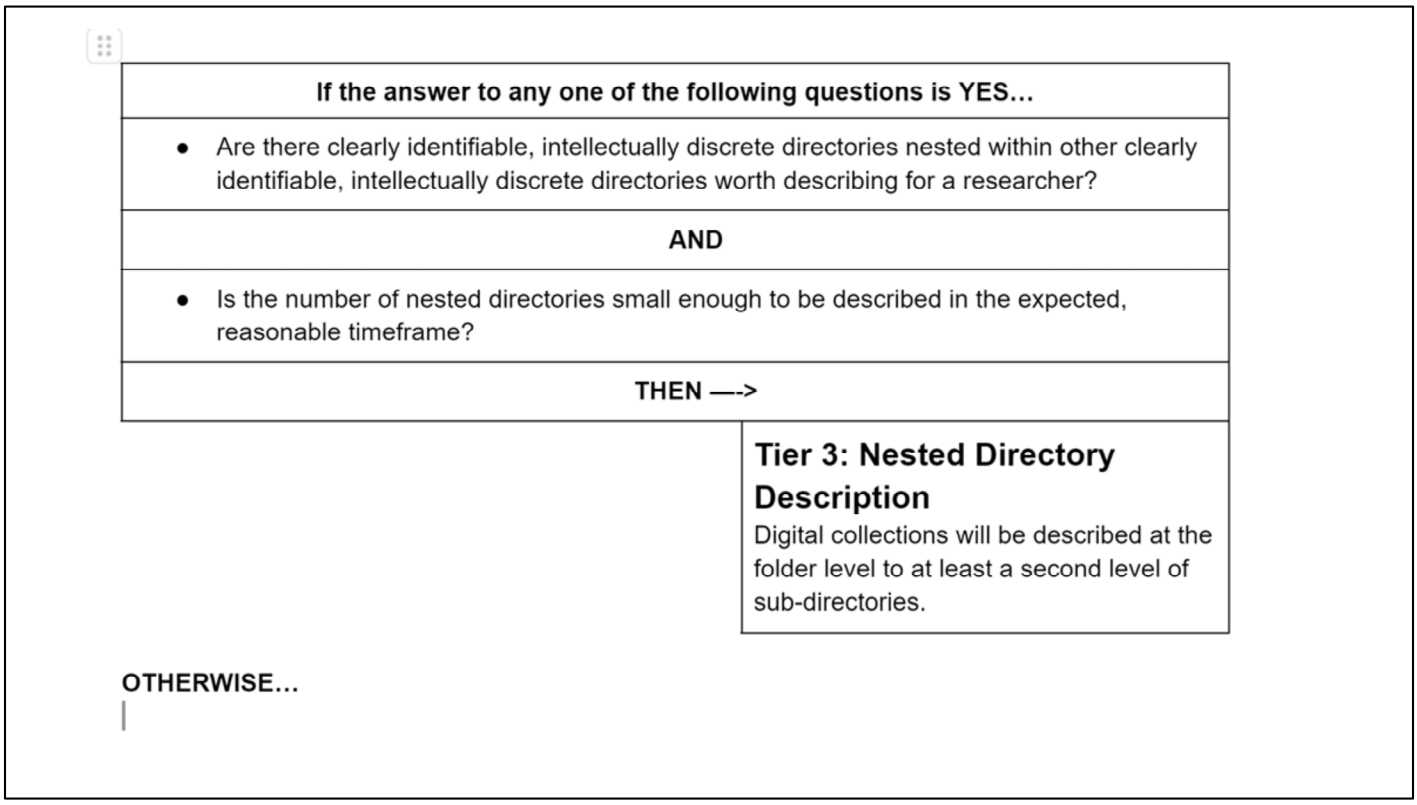
You can see we have a basic If:Then list of conditions for item level description, plus the necessary cop-out, which we have at every level. Do we have the time and resources to process at this level at this time. If both are satisfied, then yes, in this case: Item level description.

If the answer to any one of the following questions is YES…
• Are there large or more valuable digital objects that stand out from the rest of the collection, accompanied by a larger body of less valuable or distinct digital objects?
AND
• Do we currently have the time and resources to process the most valuable objects to item level?
THEN —->

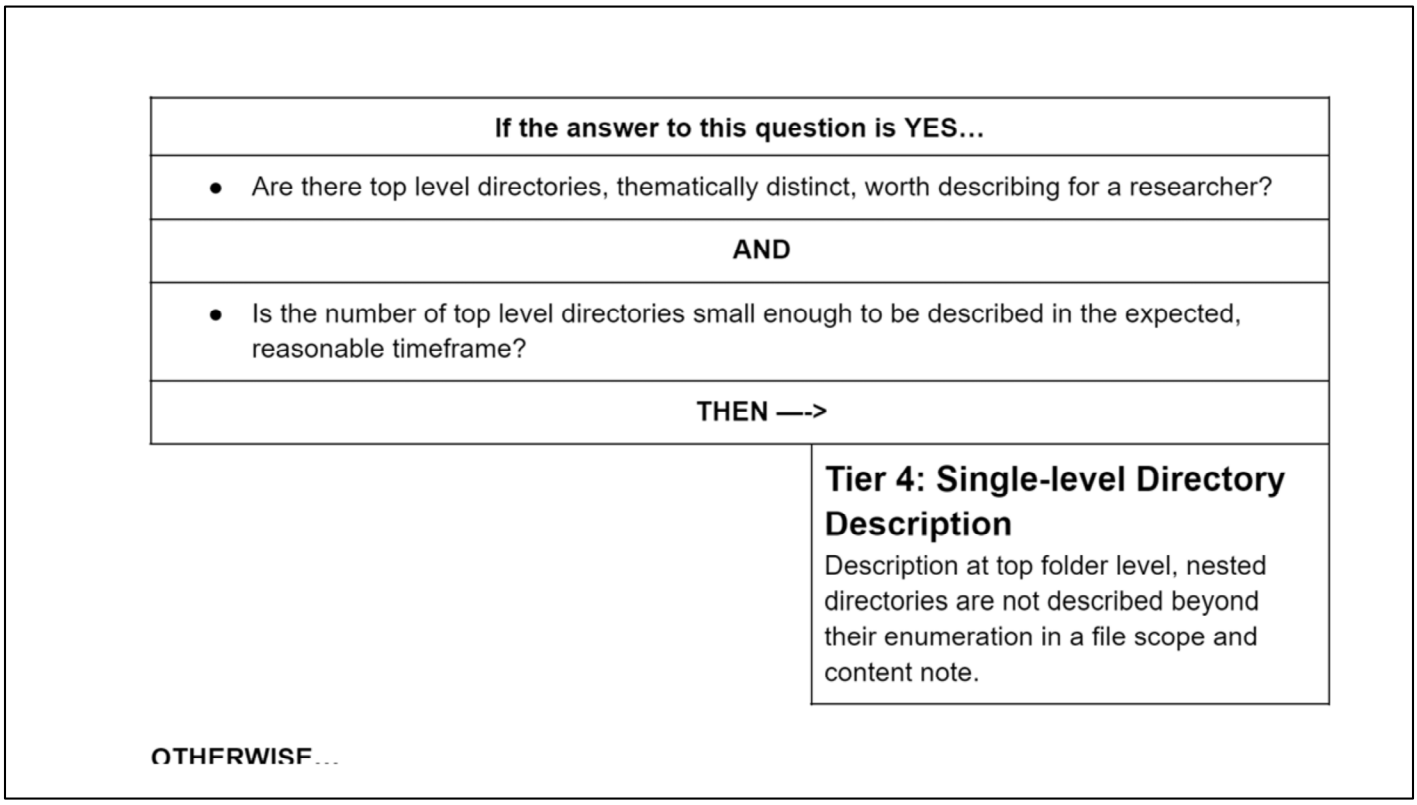
Tier 2: Partial Item + Directory Description

In these collections, some materials will be described at item level and others will receive directory, and sometimes nested sub-directory, level description.

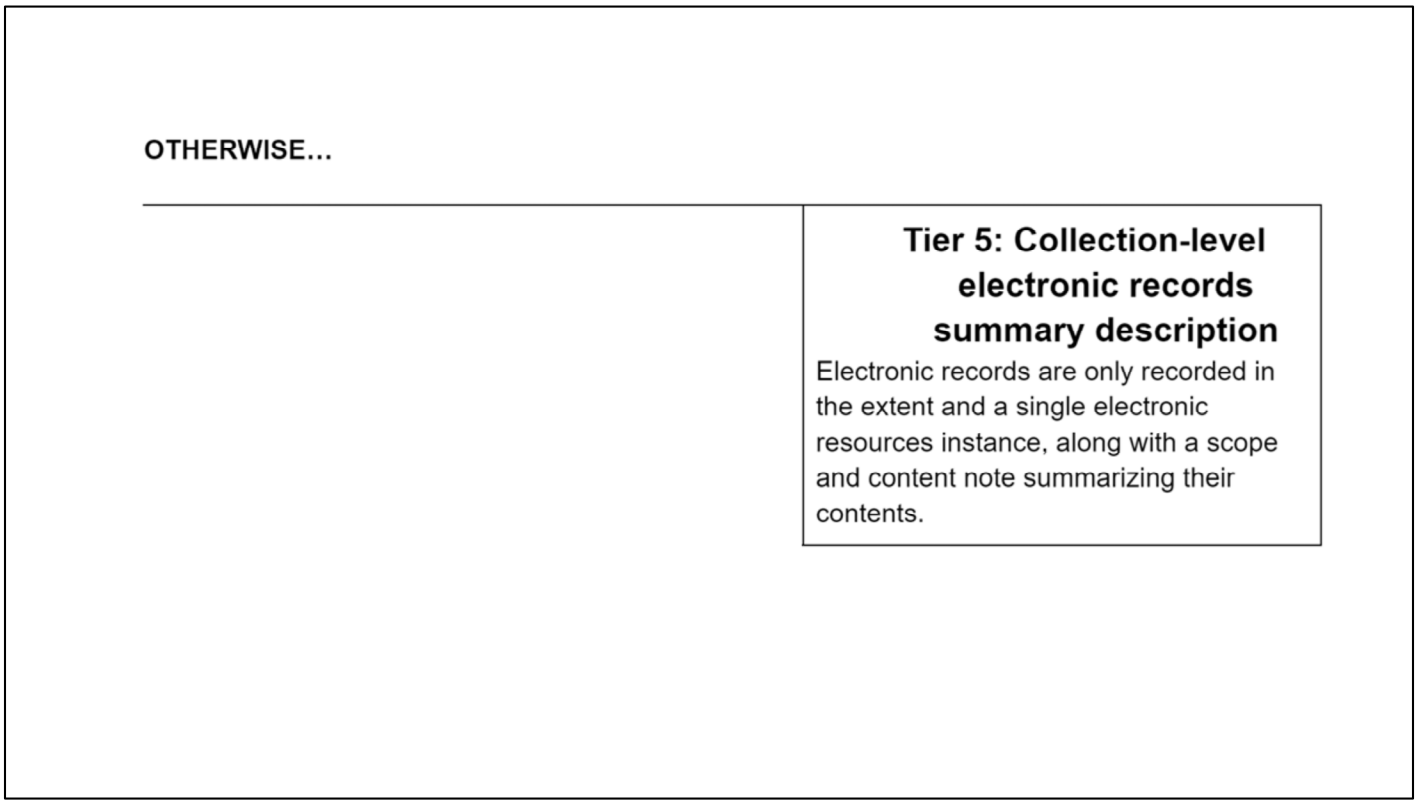
Tier 2 is similar, but composed of collections that require *some* item level description and the rest at directory level Description.This was what the Kentucky Quilt Project required.



Tier 3 is Multiple, Nested Directory Description. Digital collections in this category will be described at the folder level to at least a second level of sub-directories.

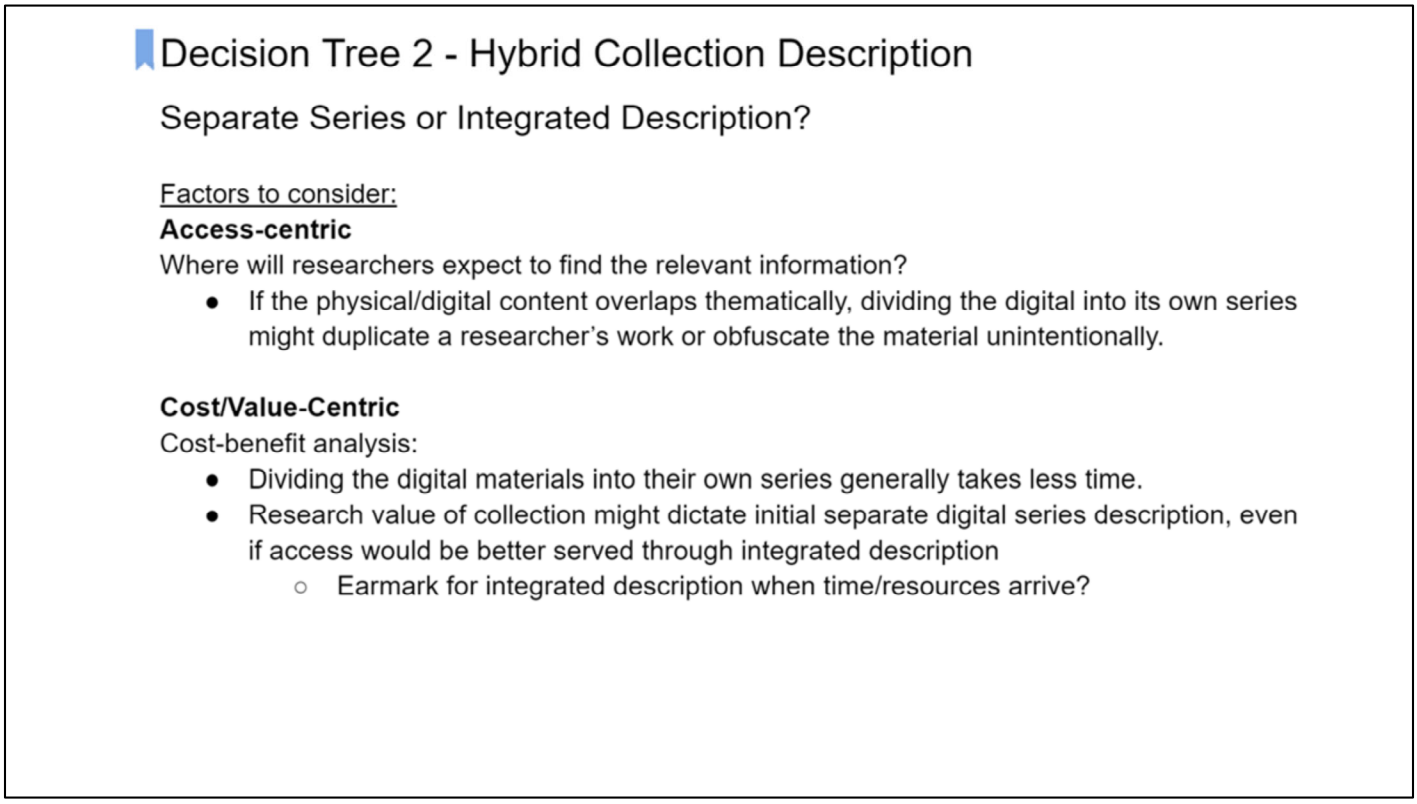


Tier 4 is Single-level Directory Description
So just the top level of directories, for collections that have a coherent top level. Directories nested beyond the first level are not described beyond their enumeration in a scope and content note, for example: 1,124 files in 36 folders.



Tier 5 is Collection-level electronic records summary description
Electronic records are only described in the extent and a single electronic resources instance, along with a scope and content note summarizing their contents. Depending on the helpfulness of such a document, this may contain a public-facing file manifest.

That level of detail is complex on its own, but considering this in light of processing the born-digital materials as part of a hybrid collection, alongside physical materials, adds another layer of factors that must be taken into account.



So we made a separate decision tree for hybrid collections.
There are two main types of Factors to consider:

Access-centric
Where will researchers expect to find the relevant information?
-If the physical/digital content overlaps thematically, dividing the digital into its own series
might duplicate a researcher's work or obfuscate the material unintentionally.

And of course, there's the Cost-benefit analyses archivists are always making:
Dividing the digital materials into their own series generally takes less time. The research value of a collection might dictate that initially we should create a separate series for digital items, even if access would be better served through integrated description.
So maybe we do that, but earmark the collection for integrated description when opportunity arrives?

Abstract Math for Archivists

IF Research Value of Collection ≥ (Time + Money to Process) THEN: Do integrated description.
IF Research Value of Collection < (Time + Money to Process) THEN: Do separate description until equation changes.

Other Factors to Consider

What *percentage* of the overall material is digital?

- If there is only a handful of digital objects in a sea of physical, taking the time to integrate the material intellectually is probably worthwhile.
- If digital and physical materials are evenly split, and the materials, while differently formatted, are intellectually alike, consider full integration or digital/physical sub-series within larger intellectual classifications.

You end with a basically impossible equation, essentially:
IF the Research Value of Collection is greater than or equal to the Time and resources to Process it, THEN: Do integrated description.
IF the Research Value of Collection is less than (Time + Money to Process) THEN: Do separate description until equation changes.

You can also consider:
What percentage of the overall material is digital?

If it's only a handful of digital objects in a sea of physical, taking the time to integrate the material intellectually seems worthwhile.

If it's evenly split, and the materials, while differently formatted, are intellectually alike, consider full integration or digital/physical sub-series within larger intellectual classifications.

Case Study 1: Evolving Description

- Fix issues with disk images
- Appraisal needed
- Decide the level of description - using the born digital decision waterfall!
- Can we easily repurpose any of the existing description?

tiny.cc/ukborndigital

Now back to the Kentucky League of Cities Records and all of its misery! With all that institutional (I want to say baggage, but a more positive term is history) – We are planning to get rid of all of those inherited disk images and the reports that have nothing in them. Conduct a hefty appraisal to decide if we need to keep some of these documents, since they were printed. Decide the level of description using our fancy born digital decision waterfall – but also see if we can repurpose any of this extremely detailed description, so as little as possible is wasted from the previous version of the description. This is part of the life-cycle of description and we don't want to throw away what's still useful. It identifies a bunch of people! So hopefully that 4.4 GB and almost 7,000 files will shrink, we will pay less and use less energy to preserve the collection – and it will be infinitely more useable…for right now.

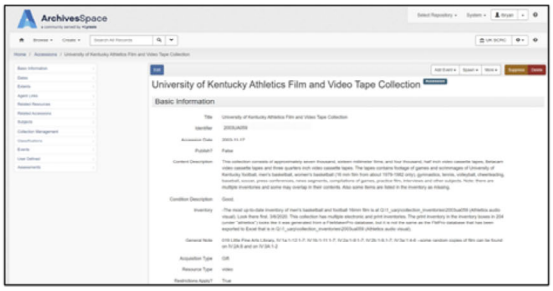
Case Study 2: Athletics Film and Videos

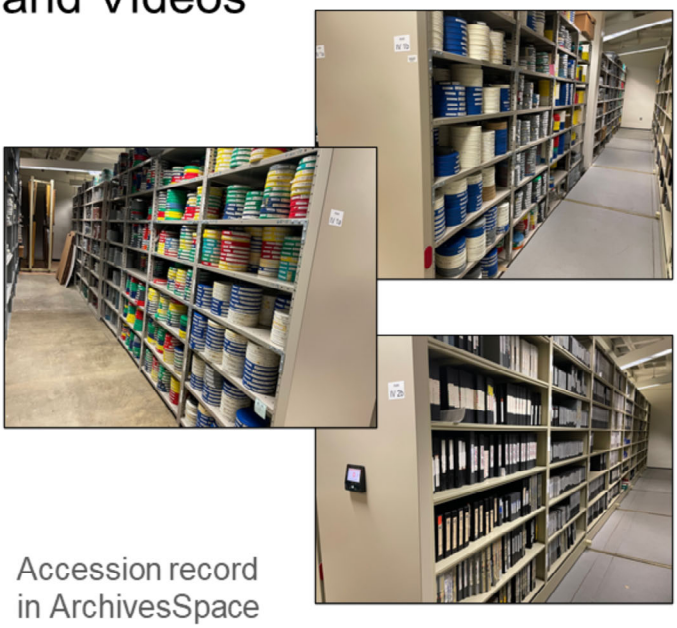
- 2003ua059: main film and video collection
- 2010ua008: specific accession of DVDs of digitized athletics content

Our second case study deals with our significant university athletics film and video collection. The core of the case study is about an accession of DVDs of digitized athletics content, their original or previous management and description, and how, why, and what we've changed or revised. But, in order to provide context for the DVDs, I'm going to first describe the main collection and how most digitized items were formerly and currently managed. Throughout, I think it'll be easier to distinguish the two collections/accessions by using their associated accession numbers as a shorthand reference. 2003ua059 (that is, the 59th University Archives accession in 2003) is the number for the main collection, while 2010ua008 is the number for the specific DVD accession.

Case Study 2: Athletics Film and Videos

- 2003ua059, the main collection
  - 16 mm film
  - Various video formats
- Football, basketball
- 1188 cubic feet
- 1933-2009





Accession record in ArchivesSpace

The main collection, that is, accession 2003ua059, comprises coaching film taken by the athletics department for football and men's and women's basketball games through about 1988, then mainly television broadcast games footage for the same sports (plus occasionally a few others) from about 1988 through about 2009. There is approximately 1188 cubic feet and the earliest film dates from 1933 while the latest video is from 2009.

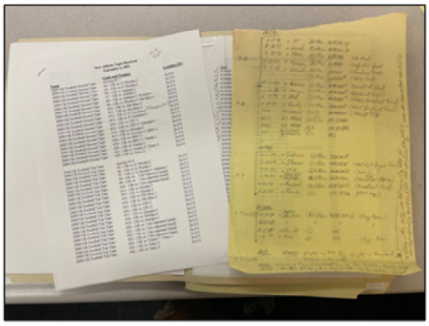
Case Study 2: Athletics Film and Videos

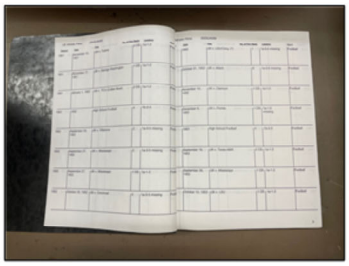
- 7” 16 mm
- 12” 16mm
- 16” 16mm
- U-matic
- Betacam SP
- DVCam
- DVPro
- Mainly arranged by sport and then by year

There are lots of formats, from various sizes of 16 mm film to U-matic, Betacam SP, DVCam and DVPro videos. The items are mainly stored by sport and then by year.

Case Study 2: Athletics Film and Videos

- Until 2014, description included:
  - Accession record
  - Older inventories
  - Season, date of game, opponent, shelf location





University athletics in general is a significant area of in-person and remote research, and many of the requests are for moving images of athletics contests. However, for years, the moving images were difficult to find without going to the shelves and looking for the physical item.

Until 2014, description of the collection included an accession record and several generations of older inventories, from typescript and handwritten to a FileMakerPro database, which was only for some of the video formats and only through 2000. The inventories always included the shelf location, season, date of game, and opponent, but the actual items were never numbered. The paper-based inventories for the film couldn't be searched, and often items that had been paged previously were returned to new locations. So, finding a particular film or video usually required going to the stacks to look for it. If a documentary filmmaker asked if we had, say, 20 different games (which happens frequently), this was a burden.

Case Study 2: Athletics Film and Videos: Evolving Description

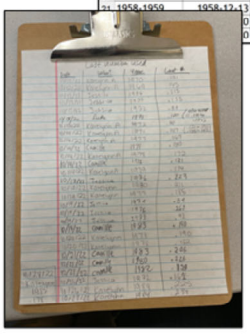
- What were they thinking! We need item numbers!

I was hired in 2011, and for several years, I wondered, “What were they thinking! We need item numbers!” I kept looking for an opportunity to start a new inventory process that would finally itemize each reel and video associated with each game.

Case Study 2: Athletics Film and Videos

- 2014: Started new inventory
- Student interns or assistants
- 6,266 reels so far
- Item number:
  - 2003ua059FB6566.01

	Season	Date of game	Item Number	Teams	Reel number	Film label	Location (IV:1B)	Digitized
2	1973	1973	2003ua059MBB7374.73	Ohio State vs. Stanford	1 of 2		10:07	
3	1973	1973	2003ua059MBB7374.74	Ohio State vs. Stanford	2 of 2		10:07	
4	1955-1956	1955	2003ua059MBB5556.03	UK vs Vandy	1 of 4		10:08	
5	1955-1956	1955	2003ua059MBB5556.04	UK vs Vandy	2 of 4		10:08	
6	1955-1956	1955	2003ua059MBB5556.05	UK vs Vandy	3 of 4		10:08	
7	1955-1956	1955	2003ua059MBB5556.07	UK Frosh vs Tennessee	1 of 3		10:06	
8	1955-1956	1955	2003ua059MBB5556.08	UK Frosh vs Tennessee	2 of 3		10:06	
9	1955-1956	1955	2003ua059MBB5556.09	UK Frosh vs Tennessee	3 of 3		10:06	
10	1955-1956	1955-11-19	2003ua059MBB5556.10	UK vs Tennessee	1 of 2		10:06	
11	1955-1956	1955-11-19	2003ua059MBB5556.11	UK vs Tennessee	2 of 2		10:06	
12	1955-1956	1956-02-01	2003ua059MBB5556.01	UK vs Duke	1 of 2	versions include	10:07	yes
13	1955-1956	1956-02-01	2003ua059MBB5556.02	UK vs Duke	2 of 2	versions include	10:07	yes
14	1955-1956	1955-1956	2003ua059MBB5556.06	UK vs Vandy	4 of 4		10:08	
15	1957-1958	1957	2003ua059MBB5758.01	Minnesota vs North Carolina State	1 of 1		10:06	
16	1957-1958	1958	2003ua059MBB5758.02	West Virginia vs Oklahoma	1 of 1		10:06	
17	1957-1958	1958	2003ua059MBB5758.03	Indiana vs Notre Dame	1 of 1		10:06	
18	1958-1959	1958-12-11	2003ua059MBB5859.02	UK vs Southern Methodist	1 of 1		8:02	
19	1958-1959	1958-12-13	2003ua059MBB5859.03	UK vs St Louis	1 of 3		8:02	
20	1958-1959	1958-12-13	2003ua059MBB5859.04	UK vs St Louis	2 of 3		8:02	
21	1958-1959	1958-12-13	2003ua059MBB5859.05	UK vs St Louis	3 of 3		8:02	
		0	2003ua059MBB5859.06	UK vs West Virginia	1 of 2		8:02	
		0	2003ua059MBB5859.07	UK vs West Virginia	2 of 2		8:02	



In 2014, I saw an opportunity with the new undergraduate diversity scholar internship to start an inventory of the 16mm film. We started with the film because the date range includes the most requested games and because UK owns the copyright, so we can give permission for use. During their year in the Libraries, a cohort of 1-3 interns rotate through individual units doing specific projects in order to get experience with cultural heritage work (among other aspects of the internship). Because the project was routine, required a small amount of training, and worked with a special format (i.e. 16mm film) it was a good option for a Special Collections 6-8 week project. Athletics is also a rich area in which to explore the constructs of race and the actions of racism, so conducting this inventory, even though it can be boring, supports the larger focus of the internship.

Through 2022, 18 Diversity intern cohorts and about 3 other student assistants have worked on this project. 6266 reels have been inventoried. Each game usually comprises two to four reels, and some are dups. The film inventory is close to completion except for the early, large football films and most of the women's basketball. (Football: 4660; MBB: 1571; WBB: 35)

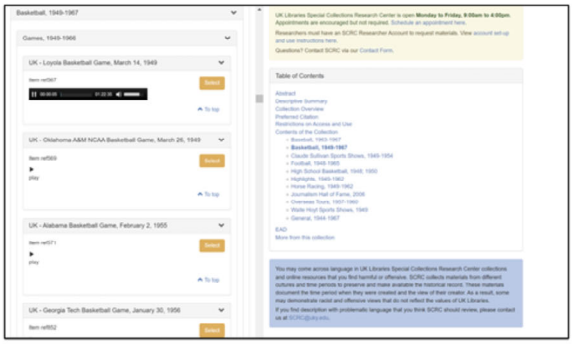
The inventory is done in Excel or Google sheets. Each physical item is labeled with an item number and then that item number is added to the new inventory along with all the other metadata elements that had been gathered before.

The item number is formed using the accession number, a letter code for football, men's basketball, or women's basketball, numbers for the season, and a consecutive number for all the reels in a season. This is example is for the first reel we inventoried of football coaching film for season 1965-1966.

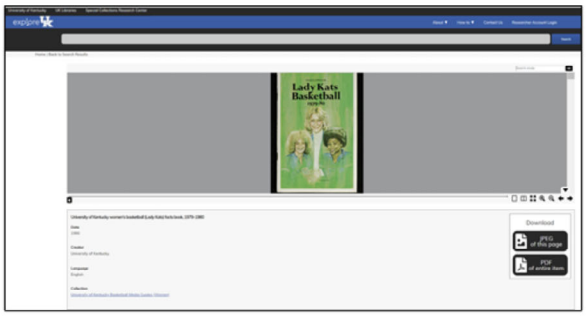
Just a note here that our implementation of Aeon in 2017 has helped with returning items to their correct locations, so that, too, makes it easier to find games and reels.

Case Study 2: Athletics Film and Videos

- Digitization by vendor on patron and athletics request
- Digitized files need to be managed and described as if they were born-digital.
- Because currently no EAD or DC as in these examples:



Claude Sullivan (athletic) audio recordings collection guide/EAD with digitized audio



Women's basketball (Lady Kats) facts book, 1979-1980 (METS/DC)

Before and during inventorying of 2003ua059, we continued to routinely digitize selected film or video on patron or athletics department request. The digitization is done by an outside vendor and is either paid for by the patron or by the Libraries. Prior to the new inventory, because there were no item numbers, we couldn't match up digitized items with analog items, and we didn't know which items had already been digitized.

Also, in addition to there being no item numbers, there is no collection guide in Encoded Archival Description/EAD or individual game description in METS with DC fields. For digital objects to be uploaded to our digital library, ExploreUK, they first need one or the other. Often, when EAD or METS is present, digitized content in general doesn't get additional description, although it does goes through the digital preservation workflow and is stored on tape. Here are two examples: Claude Sullivan athletic audio recordings collection guide/EAD with digitized audio and a women's basketball fact book from 1979-1980 with METS/DC.

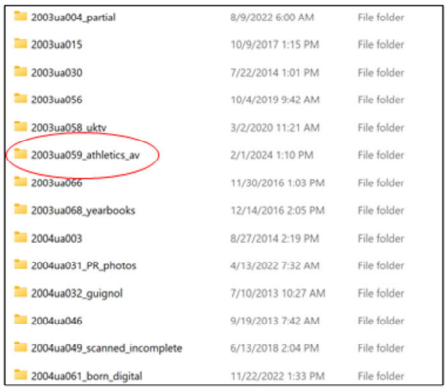
Case Study 2: Athletics Film and Videos: Evolving Description

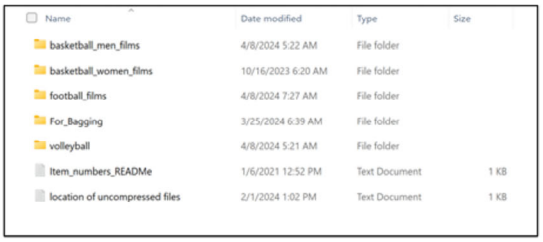
- What were they thinking! We need item numbers!
- Because no EAD or DC, digitized files need their own description, thus managed as born-digital documents.

In this case, since we're still in the stage of just creating the initial inventory, and yet we have digital objects associated with the collection because of digitization, we need to manage and describe the digitized files as if they were born-digital, equally a part of 2003ua059 along with the analog portion.

Case Study 2: Athletics Film and Videos

- Digitized items stored on a Libraries server
  - In directory under accession number (2003ua059), then in folders by sport

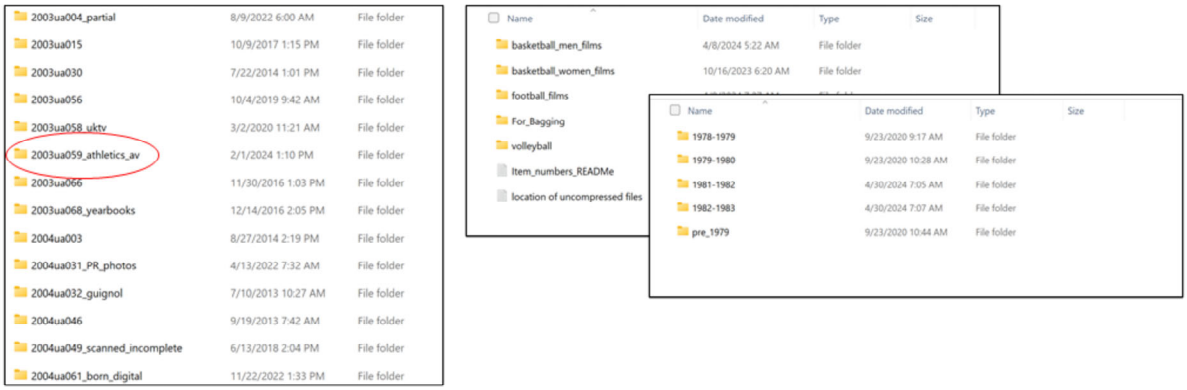




Digitized items are stored on a Libraries server. Prior to the current inventory, the only metadata available for the digital files was the associated folder and file names on the library server. First, they are stored in a directory under the accession number, then in folders by sport,

Case Study 2: Athletics Film and Videos

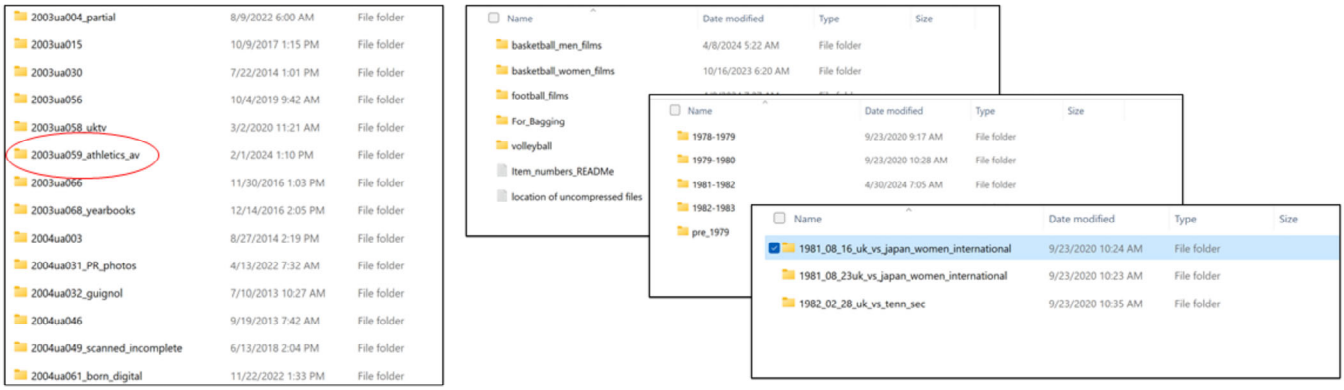
- Digitized items stored on a Libraries server
  - In directory under accession number (2003ua059), then in folders by sport, then in sub-folders by season,



Then in sub-folders by season,

Case Study 2: Athletics Film and Videos

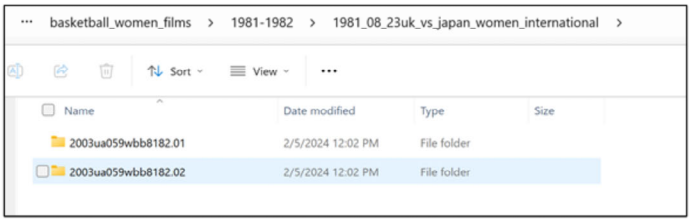
- Digitized items stored on a Libraries server
  - In directory under accession number (2003ua059), then in folders by sport, then in sub-folders by season, and further sub-folders by game date and opponent
  - Jan. 2024: moved AVI and MOV to tape and AWS



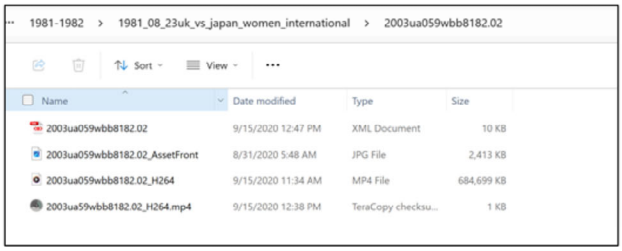
Then in further sub-folders by game date and opponent.

Case Study 2: Athletics Film and Videos

- In April of this year, starting to match digitized items to analog items
- Have more staffing available
- Locate items/reels in inventory, create sub-folder for item number, then add item numbers to files on server



2003ua59wbb8182.01 and 2003ua059wbb8182.02



2003ua059wbb8182.02: 4 files

Now that the inventory is almost complete, just in April this year, we have started to match digitized items stored on the server with their analog counterparts.

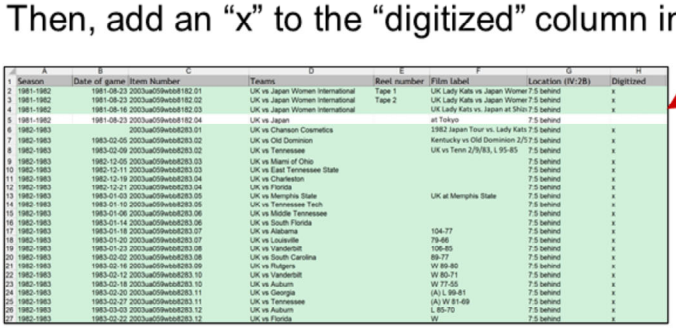
We are able to do this not only because the inventory of the analog items is mostly complete, so there are item numbers to use to connect the physical and digital versions, but also because we finally have an assistant university archivist as well as a new administrative assistant. So, we have full-time and professional staff available which we didn't have before to put concentrated time into managing the digital files.

What we're doing is searching the inventory for the game, then finding the item numbers for the relevant reels, then putting all the associated digital files into individual subfolders for each reel number. The example on the left shows two subfolders for two reels of a women's basketball game from the 1981-1982 season. The right-hand example shows the associated files for the second reel.

We also update any game description that might be missing from either the inventory that is available from the digitized copy or from the folder names that might be in the inventory already.

Case Study 2: Athletics Film and Videos

- Then, add an “x” to the “digitized” column in the inventory



	A	B	C	D	E	F	G	H
1	Season	Date of game	Item Number	Teams	Reel number	Film label	Location (IV:2B)	Digitized
2	1981-1982	1981-08-23	2003ua059wbb8182.01	UK vs Japan Women International	Tape 1	UK Lady Kats vs Japan Wome	7.5 behind	x
3	1981-1982	1981-08-23	2003ua059wbb8182.02	UK vs Japan Women International	Tape 2	UK Lady Kats vs Japan Wome	7.5 behind	x
4	1981-1982	1981-08-16	2003ua059wbb8182.03	UK vs Japan Women International		UK Lady Kats vs. Japan at Shin	7.5 behind	x
5	1981-1982	1981-08-23	2003ua059wbb8182.04	UK vs Japan		at Tokyo	7.5 behind	
6	1982-1983		2003ua059wbb8283.01	UK vs Chanson Cosmetics		1982 Japan Tour vs. Lady Kats	7.5 behind	x
7	1982-1983	1983-02-05	2003ua059wbb8283.02	UK vs Old Dominion		Kentucky vs Old Dominion 2/5	7.5 behind	x
8	1982-1983	1983-02-09	2003ua059wbb8283.02	UK vs Tennessee		UK vs Tenn 2/9/83, L 95-85	7.5 behind	x
9	1982-1983	1982-12-05	2003ua059wbb8283.03	UK vs Miami of Ohio			7.5 behind	x
10	1982-1983	1982-12-11	2003ua059wbb8283.03	UK vs East Tennessee State			7.5 behind	x
11	1982-1983	1982-12-19	2003ua059wbb8283.04	UK vs Charleston			7.5 behind	x
12	1982-1983	1982-12-21	2003ua059wbb8283.04	UK vs Florida			7.5 behind	x
13	1982-1983	1983-01-03	2003ua059wbb8283.05	UK vs Memphis State		UK at Memphis State	7.5 behind	x
14	1982-1983	1983-01-10	2003ua059wbb8283.05	UK vs Tennessee Tech			7.5 behind	x
15	1982-1983	1983-01-06	2003ua059wbb8283.06	UK vs Middle Tennessee			7.5 behind	x
16	1982-1983	1983-01-14	2003ua059wbb8283.06	UK vs South Florida			7.5 behind	x
17	1982-1983	1983-01-18	2003ua059wbb8283.07	UK vs Alabama		104-77	7.5 behind	x
18	1982-1983	1983-01-20	2003ua059wbb8283.07	UK vs Louisville		79-66	7.5 behind	x
19	1982-1983	1983-01-23	2003ua059wbb8283.08	UK vs Vanderbilt		106-85	7.5 behind	x
20	1982-1983	1983-02-02	2003ua059wbb8283.08	UK vs South Carolina		89-77	7.5 behind	x
21	1982-1983	1983-02-16	2003ua059wbb8283.09	UK vs Rutgers		W 99-80	7.5 behind	x
22	1982-1983	1983-02-12	2003ua059wbb8283.10	UK vs Vanderbilt		W 80-71	7.5 behind	x
23	1982-1983	1983-02-18	2003ua059wbb8283.10	UK vs Auburn		W 77-55	7.5 behind	x
24	1982-1983	1983-02-20	2003ua059wbb8283.11	UK vs Georgia		(A) L 99-81	7.5 behind	x
25	1982-1983	1983-02-27	2003ua059wbb8283.11	UK vs Tennessee		(A) W 81-69	7.5 behind	x
26	1982-1983	1983-03-03	2003ua059wbb8283.12	UK vs Auburn		L 85-70	7.5 behind	x
27	1982-1983	1983-02-22	2003ua059wbb8283.12	UK vs Florida		W	7.5 behind	x

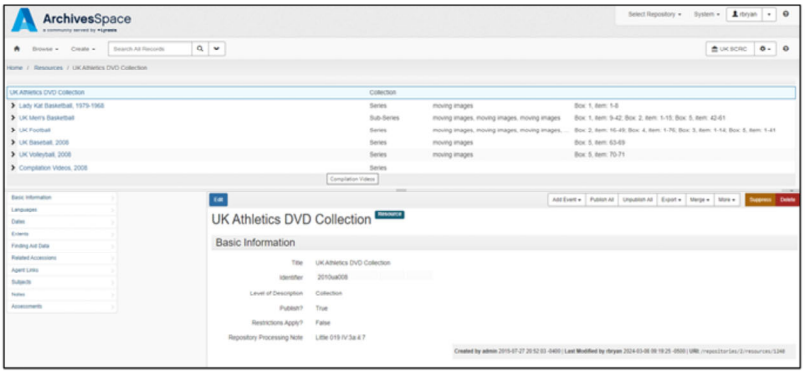
- Digitized files don’t always correspond to individual reels

Then, we add an “x” to the “digitized” column in the inventory.

A challenge is that the digitized items don’t always correspond to the individual reels, and for each case, we find a workaround for adding item numbers and making notes in the inventory.

Case Study 2: Athletics Film and Videos

- 2010ua008: 249 DVDs of digitized athletics moving images
  - Original analog items are part of 2003ua059



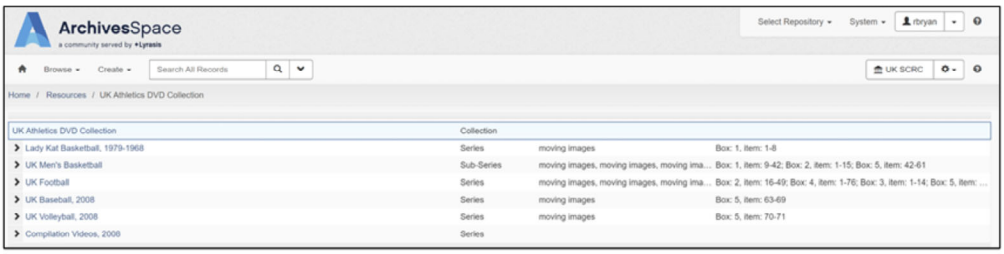
Now that I've explained a bit about the main collection of athletics moving images, that is, number 2003ua059, I'm going to talk about 2010ua008, a group of 249 DVDs of digitized film and video. The original analog items are part of the main collection. The accession came to light because of a recent patron request for the footage from a specific game. Andrew found it because the games on the DVDs had been inventoried and the inventory was in ArchivesSpace.

The accession record is bare bones. In fact, the accession description doesn't indicate that the DVD contents are digitized items from the main collection. It also doesn't have any information in it about where and how the DVDs were made, but after working with the files it appears some of the digitized games were done in-house by athletics and some by a vendor.

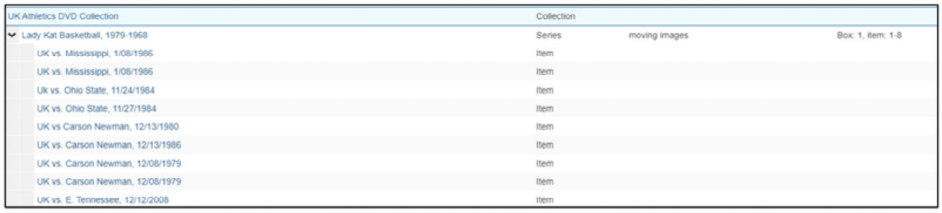
- 159 = football
- 70 = mbball
- 5=wmbball
- 2=volleyball
- 7=baseball
- 6=compilations
  - 249 discs

Case Study 2: Athletics Film and Videos

- Existing inventory is arranged by sport



- Each game is listed by opponent and date



The existing inventory is arranged by sport. The individual DVDs and the DVD boxes are numbered. The box numbers and item numbers are listed next to each sport.

Within each sport, each game is listed by opponent and date. As with the original description of 2003ua059, there is no link to the original analog item, because those items didn't have item numbers. We also didn't know if the files on the DVDs had been migrated to the server already.

Based on the game dates, which include games played after 1988, the accession includes coaching film and TV broadcasts, for which, as I've mentioned, UK does not own rights.

Case Study 2: Athletics Film and Videos: Evolving Description

- What were they thinking! We need item numbers!
- Because no EAD or DC, digitized files need their own description, thus managed as born-digital documents.
- DVDs of digitized games that had been accessioned and described as a completed separate collection and thus forgotten, need to be integrated with the main collection.

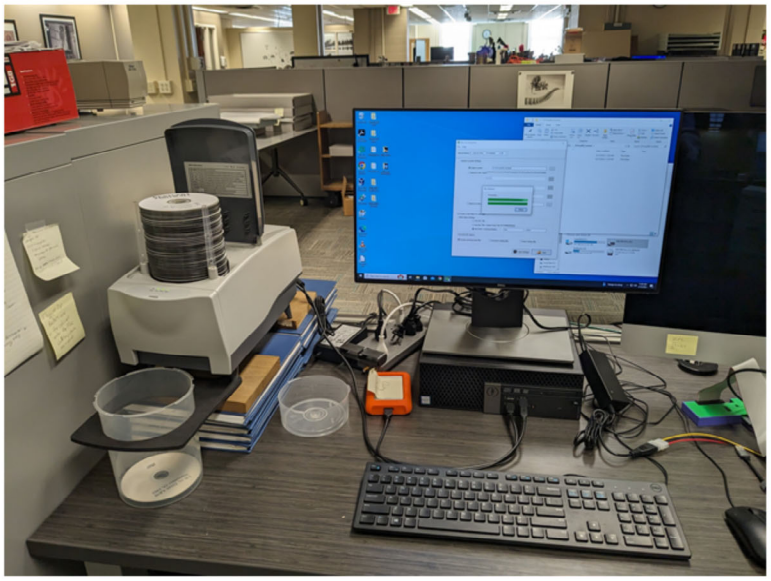
It appears that our predecessors thought that because the digitized games were stored on DVD; because in 2010 there was no easy way to handle so many DVDs; and because there was no management practice of linking digitized files with their analog originals, the DVDs needed to be managed as a separate collection. However, this led to this valuable accession from a research perspective being forgotten. Because of the high use of athletics game footage, because we now have additional staff (that is, a new Assistant University Archivist and a new Administrative Associate), we decided to explore this accession and perform born-digital management and description procedures to integrate the digitized moving images from this accession into the main collection. Andrew will explain the first steps.

Digitized and Digital Athletics Moving Images

The disks in this collection were a mixed bag, metadata-wise. Some of the disks were labeled “1 of 2” for instance, but it wasn’t clear until later intervention whether it was Copy 1 of 2 of a disk, or the first half of a football game or recording. This occurred frequently enough that the most efficient means of proceeding was to rip it all, and ask questions later.

Case Study 2: Digitized Athletics Games Moving Images

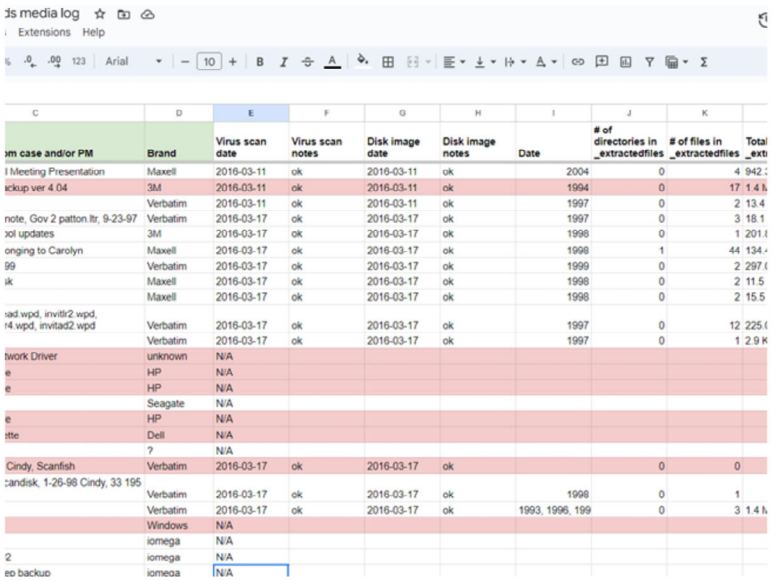
Acronova Nimbie
Disk Autoloader



Thanks to our Nimbie auto-loader, that task was largely automated.

Case Study 2: Digitized Athletics Games Moving Images

Migration tracking spreadsheet



ds media log

Extensions Help

	C	D	E	F	G	H	I	J	K
	m case and/or PM	Brand	Virus scan date	Virus scan notes	Disk image date	Disk image notes	Date	# of directories in _extractedfiles	# of files in _extractedfiles
	l Meeting Presentation	Maxell	2016-03-11	ok	2016-03-11	ok	2004	0	4
	ckup ver 4.04	3M	2016-03-11	ok	2016-03-11	ok	1994	0	17
		Verbatim	2016-03-11	ok	2016-03-11	ok	1997	0	2
	note, Gov 2 patton.ltr, 9-23-97	Verbatim	2016-03-17	ok	2016-03-17	ok	1997	0	3
	ol updates	3M	2016-03-17	ok	2016-03-17	ok	1998	0	1
	onging to Carolyn	Maxell	2016-03-17	ok	2016-03-17	ok	1998	1	44
	99	Verbatim	2016-03-17	ok	2016-03-17	ok	1999	0	2
	ik	Maxell	2016-03-17	ok	2016-03-17	ok	1998	0	2
		Maxell	2016-03-17	ok	2016-03-17	ok	1998	0	2
	ad.wpd, invitlr2.wpd, 4.wpd, invitad2.wpd	Verbatim	2016-03-17	ok	2016-03-17	ok	1997	0	12
		Verbatim	2016-03-17	ok	2016-03-17	ok	1997	0	1
	twork Driver	unknown	N/A						
	e	HP	N/A						
	e	HP	N/A						
		Seagate	N/A						
	e	HP	N/A						
	tte	Dell	N/A						
		?	N/A						
	Cindy, Scanfish	Verbatim	2016-03-17	ok	2016-03-17	ok		0	0
	candisk, 1-26-98 Cindy, 33 195	Verbatim	2016-03-17	ok	2016-03-17	ok	1998	0	1
		Verbatim	2016-03-17	ok	2016-03-17	ok	1993, 1996, 199	0	3
		Windows	N/A						
		iomega	N/A						
	2	iomega	N/A						
	eo backup	iomega	N/A						

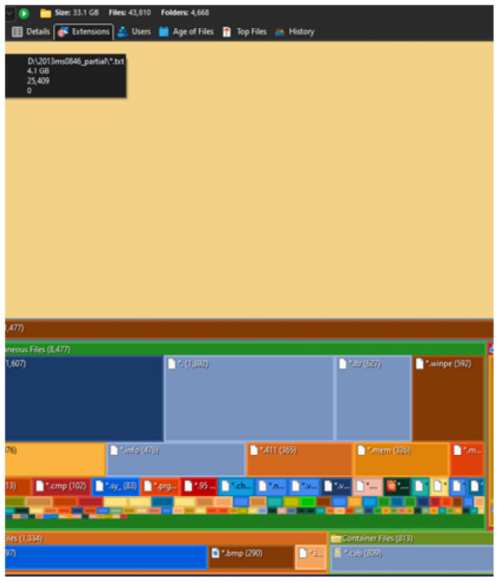
We created a spreadsheet and assigned temporary item IDs for the discs while they were being ripped.

The spreadsheet captured, wherever possible the name of the opponent or event, the date it occurred, whether a copy had already been created on our servers, the target directory name for the videos, the file size, and the number of files (some disks had 1 single 3 hour file, and some had 11 shorter standalone clips).

Case Study 2: Digitized Athletics Games Moving Images

Treesize Analysis

And Deduping

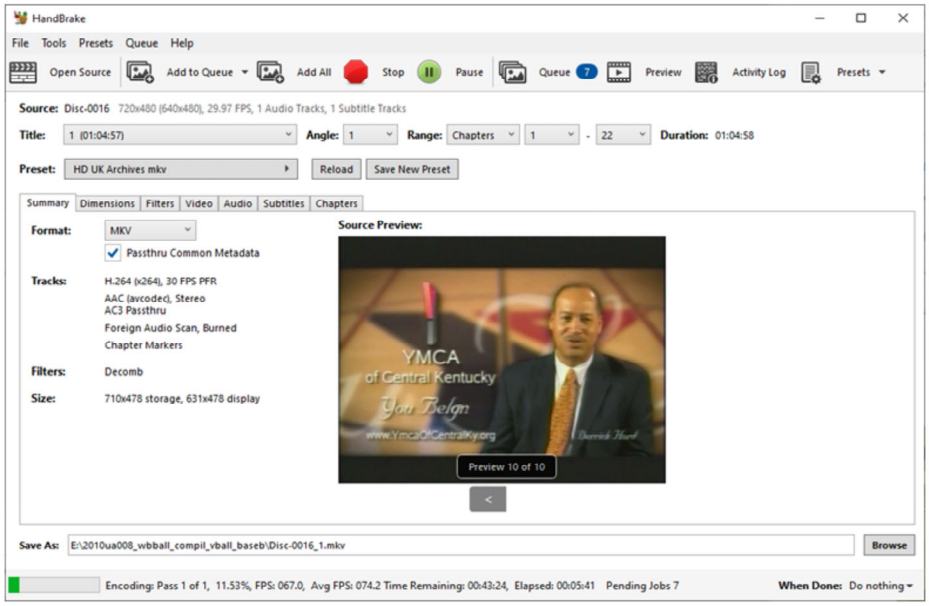


After ripping the full contents off the disks, our next automated task was to deduplicate the ripped disks. We used TreeSize to examine the completed rips and look for duplicate copies of directories.

It might seem like a lot of extra work, but ripping and deduping were both automated processes. They took time, but not my time.

Case Study 2: Digitized Athletics Games Moving Images

Handbrake conversion in absentia



Next, we converted the migrated data to standalone video files for preservation and access. I used Handbrake and built queues to convert the migrated disc data to MKV files. Again, this took only a little time up front, followed by significant background time. To be perfectly honest, I left the raw files converting while I traveled to a conference in Paris, and when I returned a week later, Handbrake was finished along with a queue identifying a handful of problematic video files. It was great. I recommend this Parisian method.

Case Study 2: Athletics Film and Videos

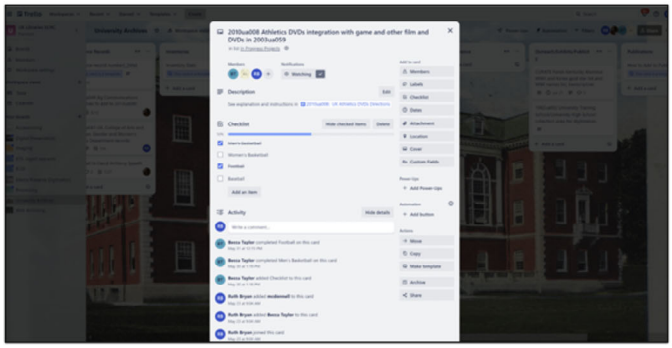
- Next description step
  - Procedures document in Google.
  - Trello card

2010ua008: UK Athletics DVDs Directions

Step 1 (5/23/2024)

Step 2

Step 3



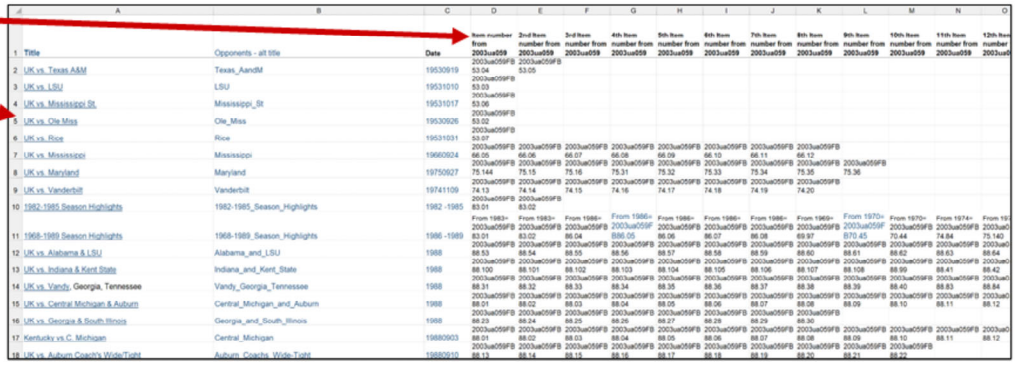
As I mentioned, we are fortunate to have several new, permanent staff who are able to work together on managing this accession. Following the digital preservation and migration steps Andrew just outlined, the next description steps can be done by the Administrative Associate during down time from the rest of her duties. I have set up a procedures document in Google docs, which is linked from a Trello card for the project in our shared University Archives projects board. This allows us to have all the information about the project in one place and to comment to each other about the project.

The description procedures identify the steps to integrate both the description and the actual digital files from the DVD accession into 2003ua059, with procedures for files that have analog counterparts and those that don't.

Case Study 2: Athletics Film and Videos

- Next description step
  - Identify the files in the migration log that are not already on the server that have inventoried analog counterparts
  - Add column in new copy of migration log for item numbers from 2003ua059

Item number columns for each game

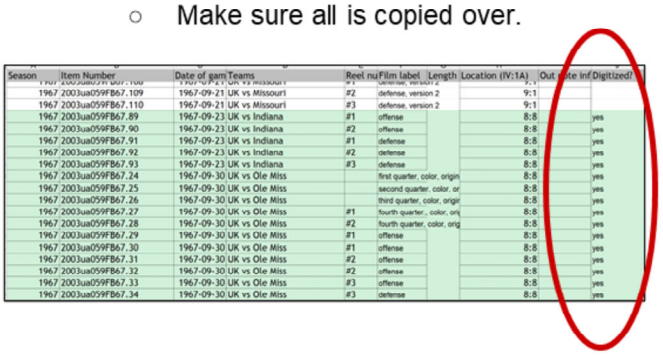


	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	Title	Opponents - alt title	Date	Item number from 2003ua059	2nd Item number from 2003ua059	3rd Item number from 2003ua059	4th Item number from 2003ua059	5th Item number from 2003ua059	6th Item number from 2003ua059	7th Item number from 2003ua059	8th Item number from 2003ua059	9th Item number from 2003ua059	10th Item number from 2003ua059	11th Item number from 2003ua059	12th Item number
2	UK vs. Texas A&M	Texas_AandM	19530919	2003ua059FB 53.04	2003ua059FB 53.05										
3	UK vs. LSU	LSU	19531010	2003ua059FB 53.03											
4	UK vs. Mississippi St.	Mississippi_St	19531017	3003ua059FB 53.06											
5	UK vs. Ole Miss	Ole_Miss	19530926	2003ua059FB 53.02											
6	UK vs. Rice	Rice	19531031	2003ua059FB 53.07											
7	UK vs. Mississippi	Mississippi	19660924	2003ua059FB 66.05	2003ua059FB 66.06	2003ua059FB 66.07	2003ua059FB 66.08	2003ua059FB 66.09	2003ua059FB 66.10	2003ua059FB 66.11	2003ua059FB 66.12				
8	UK vs. Maryland	Maryland	19750927	2003ua059FB 75.144	2003ua059FB 75.15	2003ua059FB 75.16	2003ua059FB 75.31	2003ua059FB 75.32	2003ua059FB 75.33	2003ua059FB 75.34	2003ua059FB 75.35	2003ua059FB 75.36			
9	UK vs. Vanderbilt	Vanderbilt	19741109	2003ua059FB 74.13	2003ua059FB 74.14	2003ua059FB 74.15	2003ua059FB 74.16	2003ua059FB 74.17	2003ua059FB 74.18	2003ua059FB 74.19	2003ua059FB 74.20				
10	1982-1985 Season Highlights	1982-1985_Season_Highlights	1982 -1985	2003ua059FB 83.01	2003ua059FB 83.02										
11	1968-1989 Season Highlights	1968-1989_Season_Highlights	1986 -1989	From 1983- 2003ua059FB 83.01	From 1983- 2003ua059FB 83.02	From 1986- 2003ua059FB 86.04	From 1986- 2003ua059F 886.05	From 1986- 2003ua059FB 86.06	From 1986- 2003ua059FB 86.07	From 1986- 2003ua059FB 86.08	From 1969- 2003ua059FB 69.97	From 1970- 2003ua059F B70.45	From 1970- 2003ua059FB 70.44	From 1974- 2003ua059FB 74.84	From 19 2003ua0 75.140
12	UK vs. Alabama & LSU	Alabama_and_LSU	1988	2003ua059FB 88.53	2003ua059FB 88.54	2003ua059FB 88.55	2003ua059FB 88.56	2003ua059FB 88.57	2003ua059FB 88.58	2003ua059FB 88.59	2003ua059FB 88.60	2003ua059FB 88.61	2003ua059FB 88.62	2003ua059FB 88.63	2003ua0 88.64
13	UK vs. Indiana & Kent State	Indiana_and_Kent_State	1988	2003ua059FB 88.100	2003ua059FB 88.101	2003ua059FB 88.102	2003ua059FB 88.103	2003ua059FB 88.104	2003ua059FB 88.105	2003ua059FB 88.106	2003ua059FB 88.107	2003ua059FB 88.108	2003ua059FB 88.99	2003ua059FB 88.41	2003ua0 88.42
14	UK vs. Vandy, Georgia, Tennessee	Vandy_Georgia_Tennessee	1988	2003ua059FB 88.31	2003ua059FB 88.32	2003ua059FB 88.33	2003ua059FB 88.34	2003ua059FB 88.35	2003ua059FB 88.36	2003ua059FB 88.37	2003ua059FB 88.38	2003ua059FB 88.39	2003ua059FB 88.40	2003ua059FB 88.83	2003ua0 88.84
15	UK vs. Central Michigan & Auburn	Central_Michigan_and_Auburn	1988	2003ua059FB 88.01	2003ua059FB 88.02	2003ua059FB 88.03	2003ua059FB 88.04	2003ua059FB 88.05	2003ua059FB 88.06	2003ua059FB 88.07	2003ua059FB 88.08	2003ua059FB 88.09	2003ua059FB 88.10	2003ua059FB 88.11	2003ua0 88.12
16	UK vs. Georgia & South Illinois	Georgia_and_South_Illinois	1988	2003ua059FB 88.23	2003ua059FB 88.24	2003ua059FB 88.25	2003ua059FB 88.26	2003ua059FB 88.27	2003ua059FB 88.28	2003ua059FB 88.29	2003ua059FB 88.30				
17	Kentucky vs C. Michigan	Central_Michigan	19880903	2003ua059FB 88.01	2003ua059FB 88.02	2003ua059FB 88.03	2003ua059FB 88.04	2003ua059FB 88.05	2003ua059FB 88.06	2003ua059FB 88.07	2003ua059FB 88.08	2003ua059FB 88.09	2003ua059FB 88.10	2003ua059FB 88.11	2003ua0 88.12
18	UK vs. Auburn Coach's Wide/Tight	Auburn_Coachs_Wide-Tight	19880910	2003ua059FB 88.13	2003ua059FB 88.14	2003ua059FB 88.15	2003ua059FB 88.16	2003ua059FB 88.17	2003ua059FB 88.18	2003ua059FB 88.19	2003ua059FB 88.20	2003ua059FB 88.21	2003ua059FB 88.22		

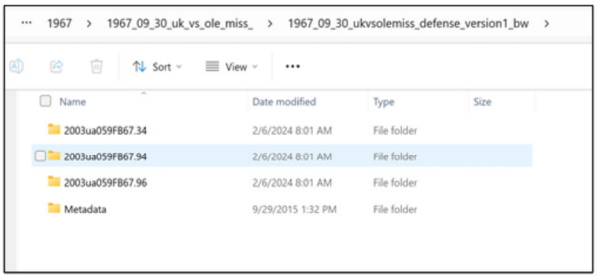
What the Administrative Associate is currently doing is using the full inventory from 2003ua059 to identify the files in the migration log that aren't already stored on the server and that have analog counterparts, and then, in a copy of the migration log, adding those item numbers to the rows of description for each file. She is also marking those games in the migration log that have discrepancies or missing or additional information.

Case Study 2: Athletics Film and Videos

- Step 2: For items with inventoried analog counterparts
  - Add “yes” to digitized column in inventory for 2003ua059
  - Arrange folders and change file names in 2010ua008 to match conventions in 2003ua059
  - Copy new folders and renamed items into 2003ua059 folder on server
  - Make sure all is copied over.



Season	Item Number	Date of gam	Teams	Reel nu	Film label	Length	Location (IV:1A)	Out	Digitized?
1967	2003ua059FB67.109	1967-09-21	UK vs Missouri	#2	defense, version 2		9:1		
1967	2003ua059FB67.110	1967-09-21	UK vs Missouri	#3	defense, version 2		9:1		
1967	2003ua059FB67.89	1967-09-23	UK vs Indiana	#1	offense		8:8		yes
1967	2003ua059FB67.90	1967-09-23	UK vs Indiana	#2	offense		8:8		yes
1967	2003ua059FB67.91	1967-09-23	UK vs Indiana	#1	defense		8:8		yes
1967	2003ua059FB67.92	1967-09-23	UK vs Indiana	#2	defense		8:8		yes
1967	2003ua059FB67.93	1967-09-23	UK vs Indiana	#3	defense		8:8		yes
1967	2003ua059FB67.24	1967-09-30	UK vs Ole Miss		first quarter, color, origin		8:8		yes
1967	2003ua059FB67.25	1967-09-30	UK vs Ole Miss		second quarter, color, or		8:8		yes
1967	2003ua059FB67.26	1967-09-30	UK vs Ole Miss		third quarter, color, origi		8:8		yes
1967	2003ua059FB67.27	1967-09-30	UK vs Ole Miss	#1	fourth quarter, color, ori		8:8		yes
1967	2003ua059FB67.28	1967-09-30	UK vs Ole Miss	#2	fourth quarter, color, orig		8:8		yes
1967	2003ua059FB67.29	1967-09-30	UK vs Ole Miss	#1	offense		8:8		yes
1967	2003ua059FB67.30	1967-09-30	UK vs Ole Miss	#1	offense		8:8		yes
1967	2003ua059FB67.31	1967-09-30	UK vs Ole Miss	#2	offense		8:8		yes
1967	2003ua059FB67.32	1967-09-30	UK vs Ole Miss	#2	offense		8:8		yes
1967	2003ua059FB67.33	1967-09-30	UK vs Ole Miss	#3	offense		8:8		yes
1967	2003ua059FB67.34	1967-09-30	UK vs Ole Miss	#3	defense		8:8		yes



football_films\1967\1967_09_30_uk_vs_ole_miss_\1967_09_30_ukvsolemiss_defense_version1_bw\2003ua059FB67.94

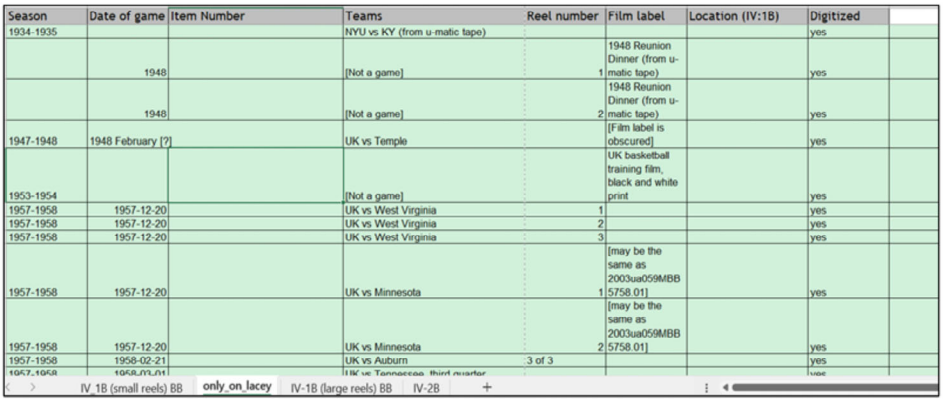
Step 2 focuses on files that have analog counterparts, which are most likely games dating from 1988 and earlier. We’ll need to add “yes” to the “digitized” column in the 2003ua059 inventory. In some cases this will require investigation, as one digital file encompasses multiple original analog items or, in some cases, we can’t immediately determine which analog items were digitized.

Then, we’ll arrange and name the folders and subfolders in the 2010ua008 folder on the server to match the naming and arranging conventions in 2003ua059, that is, by sport, then by season, then by game date and opponent, and then in folders by item number. The example on the right is for a 1967 football game that has three reels numbered 34, 94, and 95.

Finally, we’ll copy the new folders and renamed items into the folder for 2003ua059, and double check by comparing properties that all content has been copied.

Case Study 2: Digitized Athletics Games Moving Images

- Step 3: For items without inventoried analog counterparts
  - Add to “only_on_server” tab in 2003ua059 spreadsheet following existing description



Season	Date of game	Item Number	Teams	Reel number	Film label	Location (IV:1B)	Digitized
1934-1935			NYU vs KY (from u-matic tape)				yes
	1948		[Not a game]	1	1948 Reunion Dinner (from u-matic tape)		yes
	1948		[Not a game]	2	1948 Reunion Dinner (from u-matic tape)		yes
1947-1948	1948 February [?]		UK vs Temple		[Film label is obscured]		yes
1953-1954			[Not a game]		UK basketball training film, black and white print		yes
1957-1958	1957-12-20		UK vs West Virginia	1			yes
1957-1958	1957-12-20		UK vs West Virginia	2			yes
1957-1958	1957-12-20		UK vs West Virginia	3			yes
1957-1958	1957-12-20		UK vs Minnesota	1	[may be the same as 2003ua059MBB 5758.01]		yes
1957-1958	1957-12-20		UK vs Minnesota	2	[may be the same as 2003ua059MBB 5758.01]		yes
1957-1958	1958-02-21		UK vs Auburn	3 of 3			yes
1957-1958	1958-03-01		UK vs Tennessee, third quarter				yes

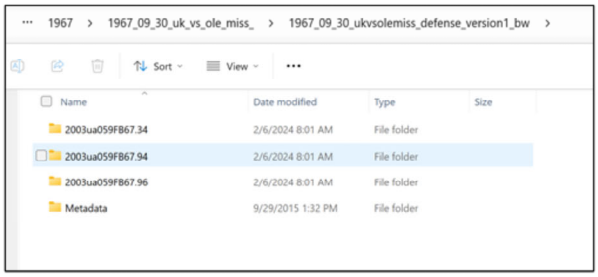
IV_1B (small reels) BB | only_on_lacey | IV-1B (large reels) BB | IV-2B | +

Season, date of game, teams, reel number, and film label (if known)

The next description steps are working with items without inventoried analog counterparts, which are most likely games from 1988 forward. The game information will need to be added to the “only on server” tab in the 2003ua059 inventory following the existing description elements, including season, date of game, teams, reel number(s), and video or film label (if known–the vendor takes a photograph of the analog item).

Case Study 2: Athletics Film and Videos

- Step 4: For items with uninventoried analog counterparts
  - Arrange folders and change file names in 2010ua008 to match conventions in 2003ua059 (will not include item number)
  - Copy new folders and renamed items into 2003ua059 folder on server.
  - Make sure all is copied over.
  - Delete 2010ua008.



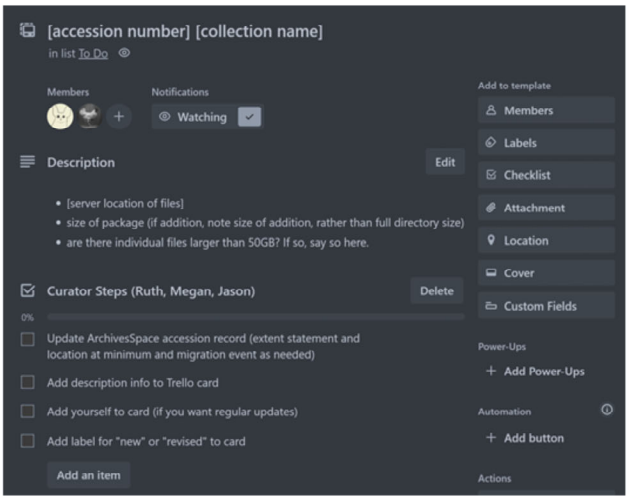
football_films\1967\1967_09_30_uk_vs_ole_miss_\1967_09_30_ukvsolemiss_defense_version1_bw\2003ua059FB67.94

Then, like for the files with inventoried analog counterparts, we'll arrange and name the folders and subfolders for the items in 2010ua008 on the server to match the naming and arranging conventions in 2003ua059, copy the new folders and renamed items into the folder on the server for 2003ua059, and double check by comparing properties that all content has been copied.

After quality checks and review, we'll then delete the 2010ua008 folder on the server. All the digital files should be in the 2003ua059 folder on the server and all the description about the files should be in the inventory for 2003ua059.

Case Study 2: Digitized Athletics Games Moving Images

- Put on Digital Preservation Trello Board
- Metadata, locations and interested parties are added to the card
- Interested parties get updates as the work progresses

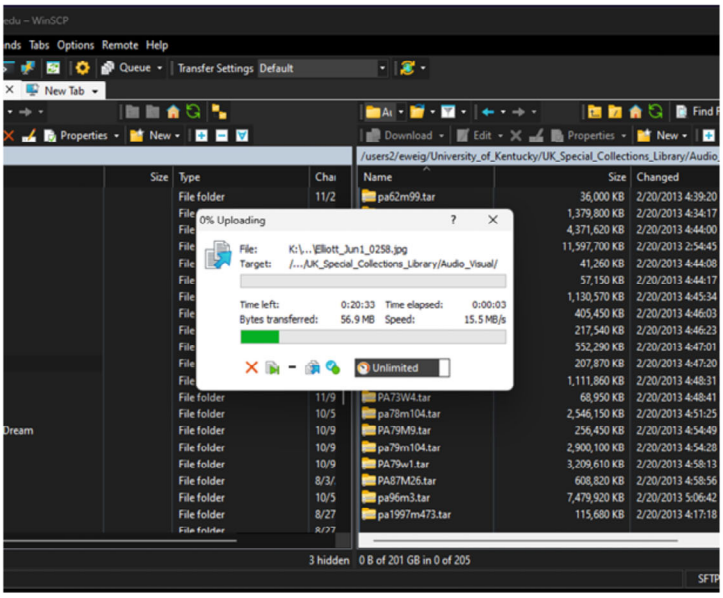


In order to ensure this wasn't all moot, we also added these materials to our Digital Preservation workflow.

This is launched when an archivist creates a new card on our Digital Preservation Trello Board and adds metadata, locations and any interested parties to get updates as the work progresses

Case Study 2: Digitized Athletics Games Moving Images

- Files are backed up to on-campus tape backup with regular fixity checks
- Also uploaded to Amazon Glacier Deep Freeze for low cost off-site backup
- Backups are tracked in a shared Google sheet

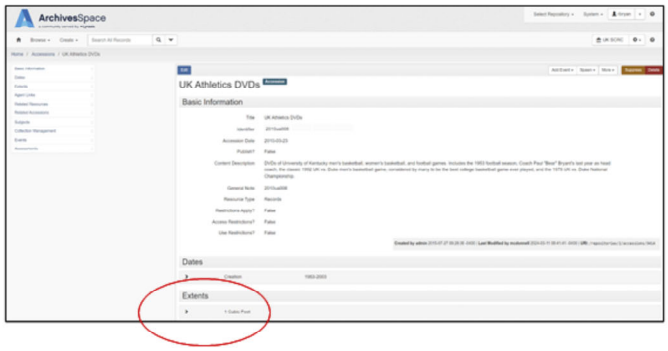


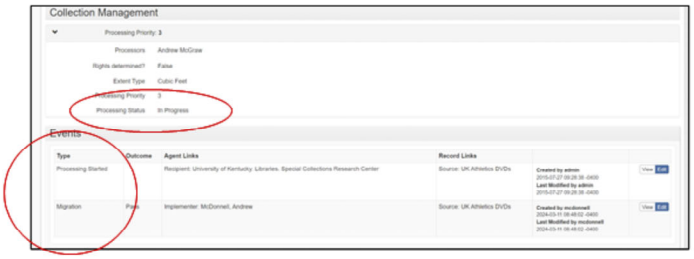
Then in addition to our networked servers (which have their own redundancy baked in by campus IT), I backed them up to on-campus tape backup which has regular fixity checks, and to AWS Glacier Deep Freeze for low cost off-site backup. I usually use the Command Line Interface but they also have a web version that has worked fine for all but the largest files I've transferred.

Along the way, I keep our backup spreadsheet up to date, and the Trello card to keep interested parties in the loop.

Case Study 2: Athletics Film and Videos

- Final description steps for DVD accession 2010ua008
  - Update accession record
  - Delete resource record



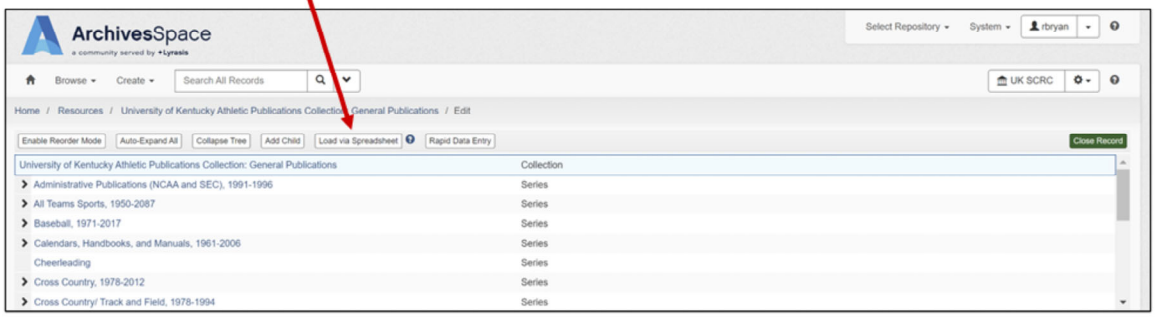


After the overall preservation steps for the main collection are complete, the final description steps include updating the accession record for 2010ua008, including extents, dates, events (including for processing and for migration) changing status to "processed," and adding a note about the location of the inventory for the main accession. The accession record for 2003ua059 should also be updated with new extent statements and dates.

The original resource record should then be deleted as it's no longer accurate or needed.

Case Study 2: Athletics Film and Videos

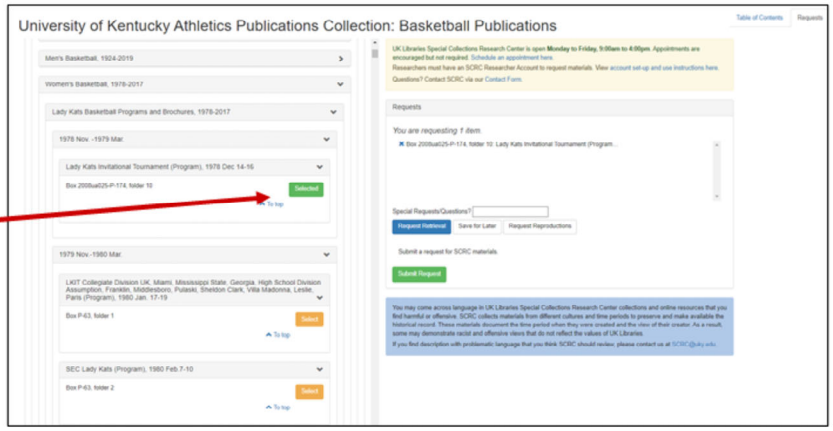
- Future steps
  - Create collection guide using inventory as base



So, whew! The description work for the athletics film and video is still just in inventory stage. We're not even at the point of making a collection guide, which is an important future step for access whether or not the digitized material is also uploaded to the digital library. We can use the Excel spreadsheet, reorganize it into import format, then import that into ArchivesSpace. Creating the import spreadsheet for a collection as large as this will take a lot of time resources, but is also a task that the Administrative Associate can do. We'll also have to create the collection guide in stages, since the video portion of the athletics moving images collection isn't inventoried.

Case Study 2: Athletics Film and Videos

- Future steps
  - Tier 1/item-level
  - Integrated description
  - Prefer to list games only
  - Upload a file manifest too?
  - Need one or more <container>s to request materials
  - Digitized content listed as born-digital content?
  - Or some or all uploaded as digitized content available for streaming?



This will be a collection guide for a high use/high value hybrid analog and digital collection with already existing item-level description that can be reused for the guide. So, that corresponds to tier 1 of our born-digital description waterfall and integrated description in our hybrid collection description tree.

However, in this case, what constitutes item level is a question. I would prefer to list games only in the collection guide rather than also listing all the reels and their associated item numbers. (The reels seem like too much information to make a collection guide readable.) A file manifest of the entire inventory could be linked from the collection guide to provide information about the reel numbers and additional reel information (such as whether its defense or offense, for example). But, since the EAD <container> tag is required in order to build the "select" button that links the collection guide to our Aeon request scripts, we'll need to figure out how to include a <container> tag for each game, when a game is comprised of 1 to 6 or 8 reels that are loose on a shelf and/or when a game also has a digitized version that is found in a digital folder. (We already have description procedures for using digital folders as <container>s.)

We'll also need to make significant management decisions that will include many of our Special Collections colleagues about whether to include some or all of the digitized content as digital archival objects available for streaming online and whether to do it using DC or EAD. This also brings up related questions about the <container> tag. If the digitized file can be streamed, should the digital file and/or the reels also be included as something a patron can request?

Case Study 2: Athletics Film and Videos: Evolving Description

- What were they thinking! We need item numbers!
- Because no EAD or DC, digitized files need their own description, thus managed as born-digital documents.
- DVDs of digitized games that had been accessioned and described as a completed separate collection and thus forgotten, need to be integrated with the main collection.
- Decisions about <container> tags for analog and digital formats of games and reels in a future collection guide.

So to recap, here are the main areas of so-called “misery” for the Athletics Film and Videos that previous archivists passed down to us that we are making changes to now based on our current expertise, knowledge, systems, tools, and available staffing resources, or the areas of “misery” we will probably be passing on to our successors.

To summarize, I determined that each reel or video needs an item number, which opens up stronger management options for the analog and digitized versions of the content and allows us to manage the digitized content like born-digital content because there’s no collection guide. The DVDs of digitized games that had been handled as a separate collection need to be integrated with the main collection. And, when that integration is complete, we move on to a new stage of collection management, building a collection guide that will rest on a foundation of our newly-created born-digital description examples and decision waterfalls or trees, but which also provides challenges to them.

The Beloved Curse:
Ever-Evolving Born-Digital Description

- It is time consuming!
- And resource-intensive
- Experimentation is good
- One size does not fit all
- Natural part of the digital lifecycle

Philip Larkin's "This Be the Verse" from Collected Poems, Farrar Straus and Giroux, 2011.
https://www.poetryfoundation.org/poems/48419/this-be-the-verse

In summary, as our two case studies demonstrate, it's time consuming and resource-intensive to keep up with and to revise and rework born-digital description and the associated policies and procedures. Our decision waterfall/decision tree is one way we are trying to codify how best to use our time and resources in the most effective way possible.

It's also important to experiment and build up established practice through specific cases and experiences.

And, finally, one size definitely doesn't fit all. Our multiple tiers of decision-making demonstrate that, as does each one of the description examples and case studies we've shared with you in this presentation.

So, we are learning to love the ever-evolving description of digital formats as a natural part of the digital lifecycle. We inherit collection management decisions and practices from our predecessors, and we pass on our best to our successors, knowing they will also shoulder this beloved curse of the constant change of technology and archival praxis.

Thank you and Questions?

Megan Mummey (she/her)
Director of Manuscript Collections
megan.mummey@uky.edu

Andrew McDonnell (he/him)
Digital Archivist
mcdonnell@uky.edu

Ruth Bryan (she/her)
University Archivist
ruth.bryan@uky.edu