

## [Slide 1]

Hello and welcome to Low-Cost, DIY, and Community-Based Approaches to Audiovisual Digitization. This is the second of two webinars in the *Strategies for Audiovisual Digitization Projects* series, which was created by the *Council on Library and Information Resources* and is graciously hosted by the *Digital Library Federation*.

I am your moderator for today, my name is Brooke Sansosti and I am the Digitization and Special Projects Lead at Carnegie Library of Pittsburgh.

Now, let me introduce our four presenters:

- Lauren Algee

Lauren Algee is Digital Curation Librarian for DC Public Libraries, Special Collections where she manages creation, preservation, and access for digital local history collections. She proposed and supervised the National Digital Stewardship Residency project that created the DCPL Memory Lab and works on the DC Punk Archive team. Lauren holds a M.I.S. from the University of Texas at Austin.

- Rachel Mattson

Rachel Mattson is an archivist and historian from Brooklyn, NY. She works as the Manager of Digital Projects in the Archives of La MaMa Experimental Theatre Club. She is also a core member of the XFR Collective and the lead organizer of the Digital Library Federation's new interest group on Government Records Transparency and Accountability. She holds a PhD in U.S. History from NYU (2004) and an MLIS from the University of Illinois, Urbana-Champaign (2014).

- Dorothea Salo

Dorothea Salo is a Faculty Associate in the Information School at the University of Wisconsin at Madison, where she teaches digital curation, metadata, database design, and linked data. She holds an MA in Library and Information Studies and another in Spanish from UW-Madison.

- Pamela Vadakan

Pamela Vadakan has managed the California Audiovisual Preservation Project since it's inception in 2010. She also teaches audiovisual preservation workshops and assesses archival collections for a variety of institutions across California for the California Preservation Program. She has a Masters in Moving Image Archiving and Preservation from New York University.

## [Slide 2]

We all know why we want to digitize our cultural heritage materials. Our motivator may be long term preservation or a focus on access, but what we may not know or have are the specialized training and associated resources needed for success.

This webinar is a step towards removing the cost and capacity barriers you may encounter in your work. Experts in the field are here to share their knowledge of low cost and DIY solutions to assist in planning your digitization and community service efforts.

**[Slide 3]**

The webinar will begin with a series of lightning talks in which the speakers will describe their work.

**[Slide 4]**

Following that, we'll have a moderated discussion, directed partly by community input; and we would like to thank everyone who filled out the online survey that was attached to the announcement.

**[Slide 5]**

Finally, we will have a Question and Answer session. Feel free to type your questions into the chat box at any point during the presentation. We'll copy them into a separate box for the presenters to answer during the Q&A time.

Let's dive right into the lightning talks. I'll pass things over first to Lauren Algee.

**[Slide 6]**

Hello, everyone. As mentioned, I am the Digital Curation Librarian at DCPL.

**[Slide 7]**

As soon as I started my position at DCPL I started hearing about public demand for digitization services; however, we wanted to do more than just provide scanners and such to the public. We decided a partnership between Special Collections and DCPL Labs (Makerspaces) would be the best strategy. DCPL Labs already does a lot of digital literacy training with the public.

We applied for and got a Digital Stewardship Resident. Jaime Mears did an incredible job with us in just under a year. You can check out her blog which is linked here. She started doing research and couldn't find any other digital libraries that provided both digital equipment and digital migration for the public. So we made that our goal. Over the year she conducted literature review, equipment selection, workflow and equipment testing, public outreach, class development, and finally launched and evaluated the lab. Let's take a look at the lab first.

**[Slide 8]**

This slide shows you all the formats we transfer. As you can see, we are very AV heavy. The lab opened at our central branch in February 2016 and it just moved to our northeast branch in April due to a renovation. It's two digitization stations that are booked through an online reservation system. Within a month of opening it was booked two months in advance. There are 3 hour sessions for AV transfer and 1 hour sessions for scanning. There are weekly 1 hour drop ins in which people can stop by and get comfortable with the equipment and ask questions. That's very important since it's a DIY lab; people need to be able to follow instructions and take themselves through very detailed workflows. The Memory Lab is part of another computer lab at the library so there isn't a lot of room for handholding. It is managed by branch staff at that library in coordination with the lab staff. The VHS and scanner are our

most used technologies. Anecdotally, our biggest users are retirees, creative folks, and stay at home parents.

**[Slide 9]**

How do we make this work? The most important things are resources. We have a libguide and “DIY” of course means that you need to have very detailed instructions. This includes all of our very detailed workflows, as well as broader information that gives people context for what they are doing (best practices, classes we offer, resources for physical preservation and digital, links to free and open source tools, notes for how to build your own lab at home).

**[Slide 10]**

There was some difficulty in weighing archival best practices against public needs in terms of file quality and encoding. We came down on really trying to build preservation into digitization at every step. That includes recommended file formats and building in metadata steps in our workflows, file naming best practices. We can only do so much before they walk out the door. And this is where our programs and classes come in.

**[Slide 11]**

We offer three classes regularly that are repeated throughout our system: Digital Preservation 101, Digital Estate Planning, and Digital Archiving and Social Media. Jaime developed and tested these classes and staff were trained to offer these.

They teach critical thinking rather than specific tools. It’s more about how to evaluate software and social media and storage platforms to meet long term needs. We also offer other events that get the public interested and involved: DC Zinefest, Home Video Day, and Preservation Workshops.

**[Slide 12]**

How do we keep the lab moving? It’s all about staffing and sustainability. Lab staff are the ones who manage appointments, troubleshoot the equipment, and teach the classes. There’s also a working group that includes special collections staff, lab staff, and now staff from the branch where the lab is located. As part of the DCPL development program we offer a class in digital preservation to get staff knowledgeable about it. When they complete it they take the knowledge back to their branch and do a program or offer a class in digital preservation. This bring the lab out of its central location and into each branch.

**[Slide 13]**

We are continuing to grow and we plan to bring parts of the lab to another DC branch before the end of the year. We also are planning to make a mobile lab truck. We are going national thanks to an IMLS National Leadership Grant.

**[Slide 14, Slide 15]**

The XFR Collective is a community-oriented, autonomous, non-profit organization that offers low-cost audiovisual digitization, preservation, and access services primarily to artists and

grassroots organizations. We also foster a horizontal mentoring model for emerging archivists, and we engage in a broad-range of education and outreach partnerships.

**[Slide 16]**

At the core of the XFR Collective's mission is lowering the barriers to audiovisual media preservation, especially with regard to material that documents the lives and work of people from marginalized communities - and especially for materials that haven't been (and may never be) acquired by a collecting institution.

**[Slide 17]**

When it was first established, in 2013, XFR Collective focused exclusively on offering very low-cost, preservation-level video digitization services for artists and small organizations. For a very small fee, partners could get six analog videotapes transferred. All transfers were performed by Collective members using the equipment had been assembled and was maintained in the home of one of the collective's members.

We had then, and still have, the capacity to transfer a diversity of video and audio formats – including Umatic, VHS, Hi-8, and Beta.

**[Slide 18]**

We developed a set of preservation-oriented workflows for our transfers using a wide range of standard calibration equipment and a combination of open-source and proprietary software platforms. We digitize to the 10-bit Uncompressed video file format, and whenever possible we upload these files to the Internet Archive.

Although this partner-based service-model still comprises the core of our work, in the past few years we have also begun to expand our mission to encompass a parallel focus on education and reproducibility.

**[Slide 19]**

We have:

- built pop-up video digitization stations and performed transfers in local community spaces;
- lent equipment to community organizations who wish to conduct in-house digitization projects;

**[Slide 20]**

- conducted workshops for grassroots groups such as the Asian American Oral History Collective who need help managing their mission-critical born-digital video and audio files;

**[Slide 21]**

And we recently entered into a partnership with the Metropolitan New York Library Council (METRO) to design and build an AV transfer station for their new educational studio. This partnership is ongoing; keep your eyes peeled for more info about it.

That's about as detailed an overview of the XFR Collective as I can give in this short presentation. But before concluding I want to highlight a few of the tools that we have found useful in doing our community-engaged work.

**[Slide 22]**

The first tool is really a set of practices and principles that we have come to call “horizontal mentorship.” The work we do requires such a broad spectrum of specialized skills and contextual knowledge – very few people, even skilled archivists, have developed all of these competencies. But each of us brings to the work a distinct body of knowledge. So we use a deliberate skill-sharing and peer-to-peer mentoring model to share our knowledge with each other.

**[Slide 23]**

Teaching each other to tinker, troubleshoot, and repair obsolete equipment is a big part of this horizontal mentorship model. When you're working with obsolete technologies like analog video equipment, so much of the knowledge about how to use and maintain that equipment is also obsolete. So we find it crucial to accumulate and share this knowledge with each other.

**[Slide 24]**

We also rely heavily on a wide range of documentation and digital tools; I've listed a few of the most handy ones below:

- AMIA Open Source Committee, Workflows for A/V Archiving <https://github.com/amiaopensource/open-workflows>
- Ashley Blewer's Minimum Viable Station document <https://docs.google.com/document/d/1oJvr8zCMK4A97GF9xYOM0uijDqyNStuwjtZ23yMRkGw/edit>
- Ethan Gates's Cable Bible <https://amiaopensource.github.io/cable-bible/>
- vRecord, open-source digital video capture software <https://github.com/amiaopensource/vrecord/blob/master/CONTRIBUTING.md>
- Texas Commission on the Arts, Videotape Identification & Assessment Guide <http://www.arts.texas.gov/wp-content/uploads/2012/04/video.pdf>
- Internet Archive (visit our landing page!): <https://archive.org/details/xfrcollective>

I'm happy to answer questions about any of these resources, or recommend additional tools, either in the Q+A or via email after this webinar ends.

**[Slide 25]**

Finally, I want to highlight a few of the resources that XFR Collective has created and make publicly available:

- XFR Workflows <https://github.com/XFRCollective/WorkflowDocumentation>
- XFR-METRO partnership zine <https://xfrcollective.wordpress.com/metro-xfr-rack-zine/>
- How to build a (preservation-oriented) A/V transfer station from scratch <https://xfrcollective.wordpress.com/2017/06/06/project-update-xfr-metro-partnership/>

· Re: Hi-8 Video

<https://xfrcollective.wordpress.com/2017/03/18/%f0%9f%91%8b%f0%9f%8f%bb-8-%f0%9f%93%bc/>

· Cataloging and File Management workshop for the Asian American Oral History Collective

[https://docs.google.com/presentation/d/16LtCx4WQYYSitfo9HxRstIVCW2T8RG0sSoqGkbrlghU/edit#slide=id.g10b18693ea\\_1\\_423](https://docs.google.com/presentation/d/16LtCx4WQYYSitfo9HxRstIVCW2T8RG0sSoqGkbrlghU/edit#slide=id.g10b18693ea_1_423)

**[Slide 26]**

More resources are available via our website. Please free to reach out to XFR Collective, or to me personally with your questions!

**[Slide 27-46]**

Hello, my name is Dorothea Salo--you can call me the mad information scientist. I'm from UW-Madison. So let me start with LIS 688: Digital Curation. We always envisioned this class as service learning course; people going out into the real world and solving real world problems, such as AV digitization and media archaeology. And we couldn't solve these problems without the equipment.

But, what if we did have the equipment? 80/20? We certainly can't tackle everything so we tackled what we could. And we successfully built it: RADD (Recovering Analog and Digital Data). This is an old picture, we actually now have a server rack.

But, we still had a problem. The information school has a thriving online program and I can't give access to the equipment to these folks. Also, we couldn't help small collections held by small organizations that didn't want to turn over all their valuable materials to be digitized off-site.

So we came up with a plan in two boxes. One is named PROUD (Portable Recovery Of Unique Data) and one is named PRAVDA (Portably Reformat Audio and Video to Digital from Analog). The idea here is to pack as much as RADD as possible into boxes that can be shipped. This allows us to get more people involved in this type of digital reformatting work.

PROUD/PRAVDA are made possible by the IMLS to whom we are infinitely grateful.

That's all I can fit into this short presentation, but I am more than happy to go over the nuts and bolts with you all!

**[Slide 47-48]**

Hi everyone. At the California Audiovisual Preservation Project we assist libraries, archives, and historical societies across California to select, digitize, and preserve archival materials related to state history. All thanks to NEH, IMLS, and the California State Library.

For the past 7 years we have focused on analog audiovisual recordings, but we have also acquired existing digital collections, all content types, and of course all sorts of files. And

starting this year we're now digitizing everything, including audiovisual recordings, photographs, scrapbooks, ledgers, and more. The result is that we've built a unique research resource, the California Light and Sound Collection.

**[Slide 49]**

We have almost 10,000 objects up from over 150 partner archives. Most of the materials selected are unique, unpublished original primary sources. Living, breathing history as it happened. All of our recordings are up at archive.org and that's where you can actually see our content, but we share our metadata with Calsphere and DPLA.

You'll see that we have a wide range of content: interviews, oral histories, speeches, debates, travelogues, etc.

**[Slide 50]**

In 2010 California Preservation Program established CAVPP to address 3 major issues: analog AV recordings are the most at-risk, most repositories have limited technical ability to take care of these recordings (and storage can be daunting with such large digital files), and folks have limited resources. The reality is that many of our partners already have a lot going on and AV collection care is not necessarily a priority.

**[Slide 51]**

So we try to make our standards minimal from start to finish as a way to encourage institutions to adopt them. The essential steps that guide us and our partners are on this slide. I will go over three of them today.

**[Slide 52]**

The first step is inventory. Have a sense of what you have. We ask partners to do this before they start nominating recordings. There is limited funding so you need to know what is the most important to digitize first.

Where are the materials? Consider value, physical condition, and threat of obsolescence. We ask partners to think about this in order to justify what needs to be preserved.

**[Slide 53]**

Describe. We ask people to start with an Excel spreadsheet and to use these minimal fields. Keep it simple. "Unknown" is a fine value to put in if you don't know at the time; you can always enhance description after digitization.

**[Slide 54]**

I'd like to give a shout out to AVCompass as a wonderful free tool. It spits out a nice spreadsheet with all your data after you've gone through the form.

**[Slide 55]**

Once we get through the nomination phase, the partner sends us the recordings. We actually outsource the digitization; we don't have the staff to transfer thousands of recordings. We batch transfer to different vendors.

We rely on the vendors to have good, calibrated equipment and knowledge of best practices.

And, of course, they provide 1:1 and batch transfer options. I do recommend outsourcing if you have a problematic format and to ask questions!

**[Slide 56]**

The last step I want to cover is quality assurance. We can't check everything 100% in real time, so we require the partners to help us with this. We check the technical specifications, the beginning/middle/end (at least 10% of the run time), the PBCore metadata created by the vendor, we run the checksum, we confirm that the quality is adequate for patron use, and we check that the file corresponds with the descriptive metadata (this part is where it helps to have the partners step in).

**[Slide 57]**

We've learned that it takes practice to learn how to differentiate analog look/sound from digital look/sound/feel and to determine if artifacts have been introduced in the transfer process. In our workshops we try to teach people about this and give examples. Of course, we always advise getting second opinions. And if you have cheap labor (student assistants, etc.) it is always great to have them step in and help.

**[Slide 58-59]**

I also advise using free tools to help automate the process. BAVC's AV Artifact Atlas is a very helpful tool when it comes to learning the look/feel and understanding where artifacts come from.

**[Slide 60]**

Thank you all so much for your presentations. Let's switch gears for our moderated discussion. During this section we will be discussing some of the challenges faced when planning and/or implementing these types of strategies. The topics are based on community feedback collected via the survey that was linked in the announcement for this series. Again, feel free to type questions into the Chat box - we will get to them during the Q&A section. We have 20 minutes for the discussion and we hope to get through all of the questions during that time.

**[Slide 61]**

First question:

What open source tools and low cost equipment would you recommend to ensure a high standard of digitization and quality control?

RACHEL: So, these presentations have been great because a lot of us are offering slides with lists of tools, many of which are open source. I would say, the tools depend on what part of the



process you are considering. For instance, for the capture process you can use AMIA open source committee's Vrecord, which is specifically designed to capture analog signal to a digital file; it wasn't created for editing videos, it was made for an archival process. It's not too hard to use and there is a lot of documentation available. I think Pamela also mentioned MD5deep, which is an open source checksum generating tool.

PAMELA: I was plugging our in house one, but it functions the same way, yes.

DOROTHEA: Small warning on Vrecord: it only works with Blackmagic capture equipment.

RACHEL: Yes, it's true that Vrecord doesn't work with all components. I would say that AMIA's Github repository is filled with tools and documentation that will help you do this work.

LAUREN: On the topic of equipment, the Data Heritage Coalition donated a lot of equipment to us. There are people out there who are enthusiasts for this type of work and they can sometimes help you out by donating, refurbishing, and repairing equipment. They can perhaps teach you a lot about how to work the equipment. Check your local vendors, too.

PAMELA: I'd like to plug the AV Artifact Atlas one more time. I'd encourage everyone to check it out because it's a good way to learn about what can go wrong in regards to digital artifacts. It also can help you learn the vocabulary that vendors will be familiar with, so it becomes easier to communicate with them.

## **[Slide 62]**

Second question:

What strategies and infrastructure do you use to provide digital storage and file management?

PAMELA: We keep our access files at the Internet Archive. I do recommend them for storage. It's free, so you can't get any better than that. Then we put out preservation masters on LTO tape. It's pretty affordable. It's about \$55 per tape, but then you'll need an LTO drive which can be around \$3,000--so it's a bit of an investment. Also the files won't be readily accessible (as if they were on a hard drive), so you must consider this ahead of time if you'll need access. It works well for us. We've had our LTO for about 7 years and we do integrity tests every year.

LAUREN: At the Memory lab we require that people bring in storage. We have a calculator tool that helps them figure out how much space they'll need for the transfer. We do recommend the Internet Archive for people who just want to put things up to the cloud and have it maintained. We also use a metadata embedding software for managing files.

RACHEL: We also use the Internet Archive. Partly because we're not a collecting institution, but partly because we work with small community-based organizations that don't necessarily have the wherewithal to install something like LTO storage infrastructure. We also are interested in helping to diversify the historical record through digitizing these resources, which requires at least some gesture towards a "preservation aware" strategy. The Internet Archive is a nice, low-

budget, accessible way to accomplish this. There are definitely issues with using the Internet Archive, it's hardly a perfect system, but it is very useful for us.

DOROTHEA: I want to put in a quick word regarding the work that IMLS and others are doing on the National Digital Platform. The storage and preservation problem is a big piece of that and they understand that there are huge gaps. Please help support that work.

**[Slide 63]**

Third question:

How do you approach and engage your institution or community with your work?

DOROTHEA: Serendipity helps. As dorky as all my acronyms are: they work, they get attention. RADD has gotten attention from local news media, for instance, because it sounds cool and interesting. I've also taken my projects to the local mini maker faire and gotten people interested in it that way. So look around. Places that don't often seem as obvious fits will surprise you.

RACHEL: XFR Collective does a variety of things for outreach. We like to bring digitization pop up stations to community spaces and other kind of more open, public spaces. We also host screenings. We have an annual (sometimes twice a year) screening in a local public garden and in universities and other spaces. We also do these workshops and maintain a good presence on the internet. Github, social media, word of mouth in our personal and professional communities.

LAUREN: I'll echo what Dorothea and Rachel said. A lot of it just has to do with getting out in the communities. It is particularly important to help people understand how more modern formats are at risk, potentially more so than the older analog formats (family videos, etc.) that they understand as more fragile.

PAMELA: We tend to work directly with organizations, rather than the public. We depend on our partners to do outreach and gather the content that we are interested in. But we've definitely done some events that have brought in the community. We do screenings and have gone down to the Archives Bazaar in Los Angeles. But really, we want the partners to do more outreach since they know their users. We'd definitely like to do more events where we partner with them, such as a scanning event in which the community provides photographs and such and we go through the digitization process with them.

**[Slide 64]**

This was a great discussion, now let's move on to the Q&A section.

Q: Regarding LTO storage, do you have a plan to extend storage beyond the inherent durability limits?

LAUREN: We have a migration plan. We started with LTO-5 and we're now making LTO-6. LTO-7 just came out and we're waiting for things to settle with that. Currently we have an LTO-6 drive, so we could still read LTO-7 tapes. We need to start planning our migration soon, but we aren't quite there yet. We don't build our LTO in-house, we get it from vendors. So we're waiting for vendors to step in and say "now we're on 7," but at the moment we're in a waiting period.

Q: How do you handle condition issues, especially with materials brought in by the community?

RACHEL: For XFR Collective, we are such a low-budget, volunteer-run shop, so we really don't have the capacity to do extensive cleaning, baking, mold remediation and such. We can't always transfer tapes that have serious condition issues. But we are starting to create a guide on how picking can pick a good vendor to help take care of these issues. Helping people feel confident in working with vendors is something we are trying to address. I will also say that it is important to know what your limits are; if you go ahead and try to take on moldy tapes, you may very well ruin a lot of your equipment.

LAUREN: I echo what Rachel has said.

PAMELA: We outsource everything and I know there are particular vendors who are enthusiastic about taking on conservation issues. Part of our selection criteria involves finding things in very poor condition (actively deteriorating, mold, etc.), so it's important that we be able to play matchmaker a bit.

Q: What about value-added service? Do you offer methods where transcription can be added to AV material for searchability, accessibility, etc.?

RACHEL: When we upload things to Internet Archive we try to add as much descriptive metadata as we can. We don't do transcriptions because we don't have that capacity, but we do our best to ensure that things are findable.

LAUREN: We've recommended Pop Up Archive to folks for transcription. It does have a free subscription level.

PAMELA: I know that some vendors do offer that as an extra service, but cost will vary. I second Lauren's recommendation of Pop Up Archive. We've added some audio recordings there and the transcriptions are pretty good at the free subscription level.

### **[Closing]**

Thank you all for attending the webinar; this concludes the Strategies for Audiovisual Digitization Projects series. I'd like to thank the presenters for sharing their expertise. I'd also like to thank CLIR for organizing this series and the Digital Library Federation for hosting. If you'd like to learn more about CLIR or DLF, we have provided links below.

Recordings of both the webinars in this series will be freely available on DLF's Digitizing Special Formats wiki (see the Links pod), as well as the event page that led you here. Transcripts and slides will also be available.