LIBRARY
LIBRARY OF CONGRESS

I\V/A\V

Informal Audio Visual Summit
September 15th, 2020
9:00-5:30 pm EST

Welcome

9:00-9:05 am EST

Katherine Zwaard, Director of Digital Strategy, Library of Congress, she/her/hers

Session 1: LC Share Out

9:05-10:30 am EST

This is on us! We will share recent developments in audiovisual process and access efforts from the American Folklife Center, National Library Service for the Blind and Print Disabled, FADGI, the National Audio Visual Conservation Center, and LC Labs.

- ***Unlocking 70 years of American public broadcasting through collaboration*** **(15 mins)**
  Rachel Curtis, National Audio Visual Conservation Center, she/her/hers
  Miranda Villesvik, WGBH, she/her/hers
- ***Crowdsourcing audio transcription: Zooniverse, dialect recordings, and you*** **(7 mins)**
  John Fenn, American Folklife Center, he/him/his
- ***FADGI updates on guidelines, formats, and software*** **(15 mins)**
  Kate Murray, FADGI, she/her/hers
- ***Building, Understanding, & Balancing Accessibility for NLS's Braille and Audio Reading Download BARD Mobile 2.0*** **(30 mins)**
  Judy Dixon, National Library Service for the Blind and Print Disabled, she/her/hers
  Wendy Stengel, Design & Development, she/her/hers
  Julia Kim, National Library Service for the Blind and Print Disabled, she/her/hers
- ***Q&A*** **(20 mins)**

Break & Networking

10:30-11:00 am EST

Session 2: Accessibility for All

11:00-12:30 pm EST

We'd like to hold this space to share-out accessibility efforts including (but not limited to) captioning, structured content, time stamping, and multi-lingual support.

- ***Accessibility in open source video digitization*** **(15 mins)**
  Ben Turkus, New York Public Library, he/him/his

  With its low cost of entry and unique combination of features, nothing else has the adaptable, omnivorous potential of the open source video digitization software vrecord. This presentation will combine a status update on the evolution of vrecord with a pitch for greater community participation and sponsorship. Recent advancements in vrecord will be situated within the larger constellation of accessibility. What is the potential role for open source software in the pursuit of greater accessibility, and how can its community orientation and inherent flexibility allow us to reshape and refine our priorities? Modeling a more inclusive, equitable approach, vrecord shows us what true collaboration can yield.

- ***YouTube Community Contributions: An impact case study*** **(7 mins)**
  Joanna White, British Film Institute, she/her/hers
  Valeria Dávila, Oregon State University Archives, she/her/hers

  On July 31st YouTube announced the discontinuation of Community Contributions. From the 28th September, 2020 they will delete all unpublished contributions and the service will cease to exist. In this lightning talk Valeria Dávila and Joanna White will discuss their use of Community Contributions to generate subtitles and closed captions for Media Area's "No Time To Wait" conference. They will share how this tool is used and its importance for improving access, and reflect on how this decision negatively impacts a growing international open-source archive community.

- ***Broadening accessibility during COVID telework*** **Panel Discussion (45 mins)**

  COVID-19 forced many archives and libraries to pivot to new workflows and projects that suit a telework environment. These workflows needed to accommodate staff in diverse ways - working irregular hours, on computers that are off the private networks, without access to software. These limitations also created opportunities to explore new or different priorities during the telework period. In this panel, archivists and librarians will introduce the creative captioning, transcription, and translation telework projects they developed to broaden accessibility to their collections during COVID-19 lockdown. They'll then join each other in a chaired conversation about the common issues they encountered, how returning to routine work might impact their efforts, and the state of accessibility in the Archives field.

  Transcription Tools for Enhanced Accessibility
  Casey Davis Kaufman, Panel Chair, WGBH/AAPB, FixIT+, she/her/hers

  Casey Davis Kaufman will give a brief presentation about transcription tools that the American Archive of Public Broadcasting (AAPB) has used to enhance the accessibility of its vast collection of historic public television and radio programs. She will discuss challenges faced with large AV collections and will share background and outcomes of the AAPB's experience using open-source transcription tools and crowdsourcing platforms for correcting speech-to-text transcripts. She will further elaborate on some areas of improvement for these tools, as well as report on current efforts being undertaken by Brandeis University's Lab for Computational Linguistics and GBH to develop and improve the speech-to-text pipeline and output along with additional open-source tools for automated metadata generation.

Building the plane as we fly: Creating an audio transcription project from scratch during the pandemic
Jonah Magar and Michael Laney, Michigan State University Libraries, he/him/his

After the pandemic closed Michigan State's campus, staff in the Kline Digital and Multimedia Center and the Vincent Voice Library began considering digital projects that could enhance our collections while providing continued employment for student workers. We didn't know what resources would be available to our students, so we aimed to minimize the computing requirements for participating in the project - deciding to use Kaltura's MediaSpace and its machine-generated captioning and caption editing tools as the basis for the project. We scaffolded this work with staff labour preparing the files for transcription and harvesting the transcripts. The results of the first phase of the project will be time-stamped transcripts of the Kalamazoo Valley Museum WKZO Transcription Disc Collection.

Piloting automated AV transcription at Wilson Library Special Collections
Erica Titkemeyer, University of North Carolina at Chapel Hill, they/them/theirs

Upon campus closure, staff across UNC Libraries rolled out a series of work from home projects for staff and students, including assignments to perform quality control on automated transcriptions of digitized audiovisual recordings from Wilson Library Special Collections. This project exposed some of the more challenging aspects of transcribing historical recordings, including handling of poor sound quality, song and speaker identification, the presence of harmful and derogatory language, and management of the overall scale.

Leveraging remote student employees to increase AV accessibility
Emily Vinson, University of Houston, she/her/hers

As the COVID-19 pandemic forced the University of Houston campus to close, I proposed a project to expand closed-captioning on our online AV collections while keeping our typically on-site student staff employed. Inspired by a transcription accuracy study I began as part of an internal library fellowship, and utilizing fellowship funding, this project was possible thanks to a partnership with the Libraries' Metadata and Digitization Services department. Working collaboratively, we've trained and managed twenty-seven student employees from across the Libraries to create high-quality closed captions for our online video collections and transcripts for oral history audio recordings, including Spanish-to-English translations.

Building community through increased accessibility
Kimberly Tarr, New York University Libraries, she/her/hers

When the pandemic forced the NYU Libraries Media Preservation labs to close in March, staff began supporting the work to caption audiovisual archival content. This project will highlight the labor involved in making digital files accessible--even when using automatic transcription software.

- ***Extensible modeling and flexible permissions: Providing advanced accessibility for audiovisual content online (7 mins)***
  Bert Lyons, AVP, he/him/his

  This lightning talk illustrates four underlying philosophies that drive Aviary's strategy for online audiovisual accessibility: 1) Provide a flexible data model that allows organizations to present audio and text in multiple languages and in multiple versions; 2) Combine content access statuses with multiple permissions layers to give organizations the flexibility to offer access to specific content, while also

supporting discoverability even when content cannot be shared; 3) Support time-based metadata to increase access and discoverability within audiovisual content ; and 4) Index all text and provide clear context to users about the relevance of search results at all times.

- ***Streamlining processing and enhancing NC State's audio visual collections*** **(7 mins)**
  Niqui O'Neil, North Carolina State University Libraries, she/her/hers

  This talk will discuss AVPD, an application for processing audio visual material. After an overview of our previous workflow for A/V conversion, we'll touch on the opportunities we found for enhancing our AV collections when building a new interface for staff and users. We'll cover the various procedures for captioning materials; dynamic adaptive streaming; and providing an interface for viewing progress and editing materials. Additionally, this presentation will touch on the viewer built for displaying the processed materials. Additionally we will talk about the opportunities the API provides for interoperability with other applications, including searching caption content.

- ***Q&A*** **(10 mins)**

Break

12:30-1:30 pm EST

Session 3: A/V Makerspace

1:30-3:00 pm EST

Presenters will share efforts to support and engage creatives, researchers, hobbyists, and the curious using cultural heritage audio and audiovisual resources.

- ***Designing for the DJ: Audio discovery for creative reuse*** **(25 mins)**
  Brian Foo, Innovator in Residence, Library of Congress, he/him/his

  Citizen DJ is a year-long experiment sponsored by LC Labs that invites the public to make new sample-based music from free-to-use audio and audiovisual materials from the Library of Congress. This session will start with a demonstration of the capabilities of Citizen DJ tools for musicians and producers. It will be followed by a look into the design process as a result of feedback from musicians, producers, and music educators who have used or plan to use Citizen DJ in their work.

- ***Audiovisual data in the cultural hackathon Coding da Vinci*** **(7 mins)**
  Ilias Kyriazis, GLAM Data Curator at Coding da Vinci / German National Library, he/him/his

  Audiovisual material may comprise valuable data, which open possibilities for new explorations and creative approaches. The German cultural hackathon Coding da Vinci gives since 2014 creative thinkers and tech-savvy participants the chance to build applications by using open GLAM data, of whom audiovisual data is a significant part. This lightning talk will present projects that have used such material and benefit from the curation of audiovisual resources.

- ***DanceDance annotation*** **(7 mins)**
  Susan L. Wiesner, Maryland Institute for Technology in the Humanities, University of Maryland, she/her/hers

Dances tell a story about the history and practices of a particular cultural group and can make the unique behaviors and patterns of a community legible by revealing the common movement patterns required for participation. Yet too often when movement is documented, the community within which dances are created is ignored or overlooked. Combining insights from knowledgeable members of the communities themselves with new technological tools, and using examples from the Alan Lomax collection, the DanceDance Annotation platform will improve the interdisciplinary study of dance and movement by creating structured data as the basis for well-curated digital resources.

- ***Our summer of TikTok: An experiment*** **(7 mins)**
  Caroline Frick, Executive Director, Texas Archive of the Moving Image, she/her/hers
  Elizabeth Hansen, Managing Director, Texas Archive of the Moving Image, she/her/hers

  For over a decade, the Texas Archive of the Moving Image has partnered with organizations and individuals across the state through a free digitization and educational program: The Texas Film Round Up. With public access central to TAMI's mission, we prioritize proactive outreach and have used social media platforms from the early days of Facebook, Tumblr, and more. This talk will focus on our Summer, 2020 goal of seeing if/how archival film and video would not only work on TikTok but how the platform might be leveraged for education and context. Join us to learn more!

- ***We save 2 film": Bringing the archive & 16mm library films to public arts education*** **(7 mins)**
  Alex Whelan, Association of Moving Image Archivists (AMIA) Film Advocacy Task Force, he/him/his

  The Association of Moving Image Archivists Film Advocacy Task Force runs a biennial workshop called "We Save 2 Film" that unites independent filmmakers and the New York Public Library's 16mm film collection to teach STEM principles and art-making to middle schoolers. Students have control over conceiving, processing, and shooting their own photochemical film. The workshop teaches students about archival principles like the longevity of physical A-V materials and offers student filmmakers a permanent home in NYPL's circulating collection. This talk will highlight the two workshop cycles to date and discuss future plans for making the curriculum scalable for other organizations.

- ***Q&A*** **(20 mins)**

Break

3:00-3:30 pm EST

Session 4: Notes from the Field

3:30-5:30 pm EST

Cultural heritage practitioners and those in aligned fields will discuss recent efforts to apply emerging technologies and/or innovative practice to sound and moving image content.

- ***Open source RAWcooked preservation workflow review at the BFI National Archive*** **(20 mins)**
  Joanna White, Collections and Information Developer, British Film Institute, she/her/hers

  The BFI National Archive recently undertook a preservation project that uses open source software RAWcooked to convert 3PB of DPX film scans into FFv1 Matroska video files. A major aspect of this project involves working alongside Media Area's Jérôme Martinez, developer of RAWcooked, to

improve the software through extensive testing. This mass digitisation project continues to reveal technical and operational differences between scanners and suppliers which in turn provides opportunities to test the limits of this software. This presentation will draw on my 'notes from the field', and the benefit working with open source projects can bring to many.

- ***DV rescue: Innovations in the preservation of digital videotape*** **Panel (15 mins)**
  Dave Rice, RiceCapades/ MediaArea, he/him/his
  Libby Savage Hopfauf, Moving Image Preservation of Puget Sound/Seattle Municipal Archives, she/her/hers

  This panel will review the DVRescue project. DV videotape formats face an exceptional obsolescence risk. Falling in-between professional expertise in file-based digital preservation and analog videotape digitization, DV tape are best preserved by migrating the data from the tape into a file rather than handling them as a video digitization event. A collaboration from MIPoPS and RiceCapades, DVRescue is funded by the NEH in order to research DV preservation and to create new tools to facilitate efficient transfer of data from tape to file. Presenters will show current models, research conclusions, and methods for troubleshooting DV capture and preservation.

- ***audioqc: speedy quality control via automation*** **(7 mins)**
  Susie Cummings, NPR Research, Archives & Data Strategy (RAD), she/her/hers
  Andrew Weaver, University of Washington Libraries, he/him/his

  This talk will focus on our collaborative development and implementation of an automated-machine based quality control (QC) tool for digitized audio, simply named audioqc. audioqc is built on a framework of open source tools, such as ffprobe, BWF MetaEdit and MediaConch, and compiles a substantial amount of information in an easy to read spreadsheet. This includes signal information, file conformance and checks for consistency of embedded metadata. In implementing audioqc at our respective institutions, we have found it a great aid in the rapid processing of large sets of digitized items and the quick identification of problem files/transfers.audioqc: speedy quality control via automation.

- ***AVKID (Audio/Video Known Issues Data)*** **(7 mins)**
  Dan Hockstein, University of North Carolina at Chapel Hill, he/him/his
  Morgan Morel, Bay Area Video Coalition, he/him/his

  AVKID (Audio Video Known Issues Data) is a new initiative to gather a corpus of "known issues" found with equipment commonly used in A/V preservation efforts. The focus is on subsiding the loss of knowledge that technicians and professionals might have that are documented informally in forums, databases, and brains across the globe.

- ***Trending archives: Reusing large-scale audiovisual collections on social media*** **(7 mins)**
  Rasa Bočytė, Netherlands Institute for Sound and Vision, she/her/hers

  Social media platforms have opened a gateway for archives to promote the wealth and breadth of their collections to audiences who are hungry for audiovisual content. The success of this depends on careful curatorial and editorial choices that get the right piece of content to the right audiences at the right time. We share insights from a three-year research project ReTV on how audiovisual archives can equip themselves with media monitoring and video analysis solutions to get support in every step of their workflows for audiovisual content curation, adaptation and publication online.

- ***Unleashing Milwaukee's history: Digital reconstructions from the WTMJ-TV collection* (7 mins)**
  Shiraz Bhathena, Digital Archivist, University of Wisconsin-Milwaukee Archives, he/him/his

  The Wisconsin Historical Society's WTMJ-TV Newsfilm collection is the largest 16mm newsfilm collection in the state of Wisconsin, measuring about 2 million feet and covering the daily events from 1950 to 1980. The majority of the collection is raw footage shot for the broadcast, but a handful of complete programs exist as well. This presentation will highlight digital reconstructions of two specials from the collection, one on gay and lesbian life in Milwaukee in 1972 and another on a day in the life of socialist mayor Frank Zeidler in 1960. Both were scanned at 2K from several sources and reassembled.

- ***Accessibility in a time of telework: Opening virtual door's to NARA's motion picture holdings* (7 mins)**
  Criss Austin, NARA, she/her/hers
  Ivy Donnell, NARA, she/her/hers

  Prior to the Covid shutdown the staff of the Motion Picture Preservation Lab at the US National Archives (NARA) digitized nearly 2,000 reels of 35mm film from the WWI collection and the USIA Library Stock Shot series. The WWI reels contain unique glimpses into military activities and the USIA reels contain beautiful images from across America, including exclusive footage from Pare Lorentz's *The River* and never-finished *Ecce Homo*. Now that these records are available online staff have tagged the reels and transcribed the production files to make the records more accessible.

- ***The Distant Viewing Toolkit* (7 mins)**
  Lauren Tilton, Co-director, Distant Viewing, she/her/hers

  The Distant TV Toolkit is a Python package designed to facilitate the computational analysis of visual culture. It contains low-level architecture for applying state-of-the-art computer vision algorithms to still and moving images. The higher-level functionality of the toolkit allows users to quickly extract semantic metadata from digitized collections. Extracted information can be visualized for search and discovery as well as aggregated and analyzed to find patterns across a corpus.

- ***The Memory Lab Network's Resources and You!* (7 mins)**
  Siobhan Hagan, DC Public Library, The Memory Lab Network, she/her/hers

  The Memory Lab Network is a project coordinated by DC Public Library (DCPL) and funded by the Institute of Museum and Library Services (IMLS) to build personal archiving programs in public libraries across the U.S. The project includes a total of fourteen library partners that receive training, mentoring, and financial support to create their own Memory Labs. The Project Manager will share our growing online resources available for those outside the network to learn how to build their own Memory Labs, and she will ask for feedback and suggestions on these resources from attendees.

- ***Q&A* (20 mins)**