AMP: An Audiovisual Metadata Platform to Support Mass Description
Outline

1. Background and context - Jon

2. Outcomes and next steps - Chris
The Challenge

- Growing AV collections
  - Digitization
  - Explosion of born-digital

- Increased expectations for access
The Challenge

- Many AV collections lack metadata
  - Discovery
  - Identification
  - Navigation
  - Rights
  - Accessibility
- Institutions lack resources for large cataloging/transcription/inventory/rights clearance projects
The Challenge: Indiana University

- MDPI: Media Digitization and Preservation Initiative
  - 280,000+ AV items; 25,000+ films
  - mdpi.iu.edu
- 80+ different units
- 20+ different physical formats
- Partnership with Memnon
- Variety of existing (or nonexisting) metadata
- Avalon Media System access platform
The Opportunity

- Mass digitization approach extended to AV
  - “Digitize first”
- Emergence and continued improvement of machine learning and other automated tools
- How can we leverage the best of automated tools and human expertise?
Existing Work in the context of AV Archives

- Application of specific machine learning tools
  - e.g. speech-to-text, named entity recognition
- “Black box” systems
  - One size fits all, brute force approach to automated metadata generation
- Customized workflows
  - e.g. MiCO Platform
Existing Work: Indiana University

- Consulting engagement with AVP in 2016 to identify metadata and rights workflows
- Phased approach
- Identification of MGMs: Metadata Generation Mechanisms
- Need for platform to support workflows, metadata warehouse
Stakeholders, User Requirements, & Personas
Existing Work: Indiana University

- **AMIA 2016:**
  - Chris Lacinak and Jon Dunn. “From Mass Digitization to Mass Description: Indiana University’s Strategy To Overcome The Next Great Challenge.”
  - go.iu.edu/1Pvj
Context: UT/HiPSTAS

- *High Performance Sound Technologies for Access and Scholarship*
- An assessment of scholarly requirements for analyzing sound
- An assessment of technological infrastructures needed to support discovery
- Preliminary tests that demonstrate the efficacy of using such tools in humanities scholarship
- Developing a freely available, open-source, API-driven application for general use
AMP: Audiovisual Metadata Platform

- Audiovisual Metadata Platform
- Planning grant from Andrew W. Mellon Foundation (July 2017 - January 2018)
  - Focus on technical architecture
- In-person workshop (September 2017)
- Deliverables:
  - White paper: go.iu.edu/ampreport
  - Draft proposal for implementation and pilot test
AMP: Audiovisual Metadata Platform

- Open source software platform to support metadata creation for AV collections
- Design and execute workflows combining automated and human steps
- Integrate multiple MGMs
  - Automated, manual
  - Local, HPC, cloud
AMP Conceptual Diagram

- Media Content
- Existing Metadata
- Workflow system
- MGM
- MGM
- MGM
- Enriched Metadata
- Target System
- Users

AMP
Core AMP Team

- Indiana University Libraries
  - Jon Dunn
  - Julie Hardesty
- University of Texas at Austin School of Information
  - Tanya Clement
- AVP
  - Adeel Ahmad
  - Chris Lacinak
  - Amy Rudersdorf
Mellon-funded workshop

3 Days
16 People
1 Tech Platform Plan
Workshop Participants

- Kristian Allen, UCLA Library
- Jon Cameron, IU Libraries
- Maria Esteva, Texas Advanced Computing Center, UT at Austin
- Mike Giarlo, Stanford University Libraries
- Brian McFee, Music & Audio Research Laboratory, NYU
- Scott Rife, Packard Campus for AV Conservation, Library of Congress
- Sadie Roosa, WGBH Media Library & Archives
- Felix Saurbier, TIB/German National Library of Science & Technology
- Brian Wheeler, IU Libraries
- Maria Whitaker, IU Libraries
Workshop Overview

- **Day one:**
  - Overview, scope, background, & goals
  - Requirements development
- **Day two:**
  - Components & candidates
  - MGMs
- **Day three:**
  - Scenarios
Next Steps
Example Workflow Scenario
Example Workflow Scenario

ffmpeg

Audio file

Video file with audio

AV Demuxer

Video file with no audio

Embedded metadata extraction
Example Workflow Scenario

Audio file → Silence detection → Music Detection → Mufin AudioID

Speech detection

Yes → Language detection → Speaker identification → Automated speech-to-text

No → Fraunhofer AV Analyzer

Yes → Verbit.ai Watson

No → Kaldi

Yes → Speech-to-text human refinement

Fraunhofer AV Analyzer

MusicBrainz Gracenote

Commercial music database query

Commercial?
Example Workflow Scenario

Language detection → Speaker identification → Automated speech-to-text → Speech-to-text human refinement → Named entity extraction → Sentiment Analysis → Personal identifying information analysis → GPS coordinate identification, dBPedia URI identification → Data Recon

Music Detection

Yes → Mufin AudioID

Yes → Commercial music database query

Commercial?

Yes → Beats per minute → Genre detection → Instrument detection

No → Commercial music match query
Example Workflow Scenario

- Video file with no audio
  - Scene detection
  - Object recognition
  - Facial recognition
  - Age analysis
  - Expression analysis

- MedialInfo
  - Embedded metadata extraction
    - GPS coordinate resolution
    - Date and time identification

- Fraunhofer AV Analyzer
  - VU Digital

- Script
Borrow, Build, Buy Considerations

- Ownership of “the machine”
- Black boxes
  - Very limited options
  - Metadata cultivation concept
  - Focal point
- Limited capabilities demonstrated
- Exploring building on top of MiCO
Anticipated Challenges & Considerations
Stay Tuned!
Thanks!

Jon Dunn
jwd@iu.edu  @jwdunn

Chris Lacinak
chris@WeAreAVP.com  @WeAreAVP

White paper
go.iu.edu/ampreport