

Metadata and the Future of Your Video in an Online World

John Hauser, Access Humboldt
The Community Media Archive

March 16, 2018 10:30am - 12:00pm

<https://cutt.ly/4tq2UTF>



What is Metadata?

- Data about data
- Descriptive information about an item
- Metadata for a video should answer the basic questions:
 - **Who? What? Where? When?**
 - What's specific to **this** video rather than **that** one?

Workshop Context

- It's not about archiving!
- It's about metadata
- It's about designing your metadata at a collection level, not an item level
- so that you're ready for the future

Key Points

- the future of video distribution is not TV, it's the internet
- you need consistently tagged video to get ready for the future
- metadata storage should be vendor neutral
- preparing for archiving/preservation can help prepare for the future

Why Metadata Matters

- Key to “Discoverability” in an online world
- Bridge between Cablecasting and new forms of video distribution
- It’s what feeds Search Engines and Relevance Ranking algorithms

Think about Metadata, Search

- think about how you want to group your video programs
- think about how users will find your video
- Don't lock up categories/genre in your website
 - include them in item's metadata

How to Organize Your Videos

- By Date (default “River of Time” metaphor)
- By Channel, Type, Genre, Category, Provenance
- By Series
- By Producer
- By Topic, Subject, Keyword, Tags

Organization of Metadata

- WebForm (One Video at a Time) focus is on one video
 - no consistency between metadata for multiple videos
- Spreadsheet (Row, Column Interface) focus is on the collection
 - view many rows/videos at a time

Webform Interface

Page Title *	PCAM Voices December Innocence Project.mpeg 4	✓
Page URL *	https://archive.org/details/PCAMVoicesDece...	✓
Description *	Add a description of the item page	✓
Subject Tags *	Add keywords, separated by commas	✓
Creator	Creator of the content	✓
Date	Date work was created/published	✓
Collection *	Nevada County Digital Media Center/NCTV	✓
Test Item	No	✓
Language	Language of the work	✓
License	No license selected	✓
More Options	Add additional metadata...	

Drag and Drop More Files Here or [+ Select files to add](#)

Name	Size	x
PCAM_Voices_December_-_Innocence_Project.mpeg4.mp4	41 MB	✗

Spreadsheet Interface

Item	Title	Description	Runtime	Subject1	Subject2
Video1					
Video2					
Video3					
Video4					

Specific Metadata Tips

- Make Title Unique
- Include Runtime
- Include Producer Name
- Use Multiple Subjects
- Include Series Tag
- Expand Acronyms for Non-Local Viewers/Searchers

Make Title Unique!

- Don't use (only) Series Name for your Title!
- This!
 - <https://www.youtube.com/user/wctvnet/videos>
- Not This!
 - <https://vimeo.com/citynewberryfl/videos/sort:date>

Include Runtime in your metadata!

- displays on the item's detail page
- so that a viewer can see how long it is without clicking on the video and waiting for the video to load
- metadata key "runtime"
- use HH:MM:SS format: 00:28:57

Include Producer/Creator Name

- metadata key “creator”
- metadata value “Carl Fairfield”
- displays on an item’s detail page twice!
- allows searching for a producer’s videos inside of an aggregated collection
- collection:access_humboldt AND creator:“Carl Fairfield”

Use Multiple “Subject” Tags

- they are what show up under “Topics”
- they are searchable drive to a group of items, not an individual item
 - don’t combine title and date in a single subject tag
- Faceted in Archive’s interface (clickable)

Include “Series” tag

- metadata key “series”
- metadata value “whatever”
- also add as an additional “subject” tag
- “series” tag is for automated searches and is “invisible”
- “subject” is for humans; visible as a topic

Add Language (if not English)

- metadata key “language”
- metadata value “Spanish”
- or use ISO 639 Language code
- Faceted in Archive’s web interface

EA, Expand Acronyms!

- if you use acronyms in your metadata, add the long form as well
- NBLCA, National Black Leadership Commission on AIDS
- KCB, Keep Cincinnati Beautiful
- GTEI, Gateway to Economic Inclusion
- Avoid Overloading Acronyms
 - CAC, Cable Advisory Commission, Citizen's Advisory Council

Metadata Makeovers

Before

- Title
- No Subject Tags

After

- Title
- Geographic and Organization
- Series
- Government Meeting?
- Name of Board or Commission
- Provenance (Youtube or Vimeo)
- Year

Metadata Recommendations

- Consider metadata from the collection perspective vs. the item level
- Use Consistent Naming in Titles (spelling & CAPITALIZATION!)
- Add Dates or Episode numbering to Title - make it unique!
- Use something other than a common word or phrase for the title
- Include data only you would know and probably have - producer
- Use a vendor neutral solution

Metadata and the Archive

- Minimal required – identifier, title, description, subject
- Anything accepted (and retrievable through `_meta.xml` file on the item's “detail” page)
- subject(s) – what shows up under “Topics” link for collection
- Identifier must be unique across 36+ million items!

Metadata and Archive - More

- Identifier naming restrictions
 - use the characters a-zA-Z0-9._-
 - no spaces, parenthesis, braces, pound signs, colons
 - length of identifier - must be greater than 4 characters, less than 99
- Playback server is more permissive

Takeaways

- Design metadata at the collection level before the item level
- Focus on Metadata Consistency
- Choose “DIY” Archiving or “Archiving as a Service” via YT or Vimeo
- Get Started!
- Easiest way to get your video archived is upload to Youtube or Vimeo and let me know your channel URL

Thanks!

- To Brewster and the Internet Archive staff
- To Sean and Access Humboldt for underwriting the CMA effort
- To the Access Centers that have contributed video to the CMA
- To Eureka City Schools & HCOE for the 1 Gb/s network connection

More Information

- These slides: <https://cutt.ly/4tq2UTF>
- Community Media Archive wiki: <https://cutt.ly/ztqMRCu>
- [Documentary video](#) of the Internet Archive
- [More photos](#) by Jason Scott of the Internet Archive
- john@accesshumboldt.net
- Go forth and Metadata!

