

Metadata and the Future of Your Video in an Online World

John Hauser, Access Humboldt
The Community Media Archive



WCM / Midwest Regional Virtual Conference
October 1, 2020

<https://cutt.ly/BtgkqoA>

Speaker Intro

- Known as “The Archive Guy”
- Founded the Community Media Archive in 2009
- Solved VOD storage or bandwidth for Access Humboldt
- Created the YouTube->Archive project in 2015
- Added more than 960,000 videos and 900 collections to the Archive

Workshop Evolution

- It used to be about “**DIY**” **archiving** of video files
- Now, it’s about **metadata** associated with video files
- Why the change?

The Future is Here!

- the future of video distribution is not Cable TV, it's the **internet**
- distribution platforms for online video have **proliferated**
- you need **consistently tagged** video to get ready for the future
- metadata design & storage should be **vendor neutral**

Making the Transition

- Search Engines work on **text, not video!**
- Access Centers produce **video**
- Centers should **supplement** their video with **text** describing and categorizing each video

What is Metadata?

- **Descriptive information** about an item
- Provides the **context** for the item
- Answer the basic questions:
 - **Who? What? Where? When?**
 - What's specific to **this** video rather than **that** one?

Why Metadata Matters

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

Think about Metadata, Search

- What are the ways you **group** your videos?
- How will users **find** your video w/o a direct link?
- Are **categories/genre/series** “locked up” in your website?
 - include them in item’s metadata, so they’re **searchable**







How to Organize Your Videos

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

Data Entry vs. 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, identify inconsistencies

Web Form Interface

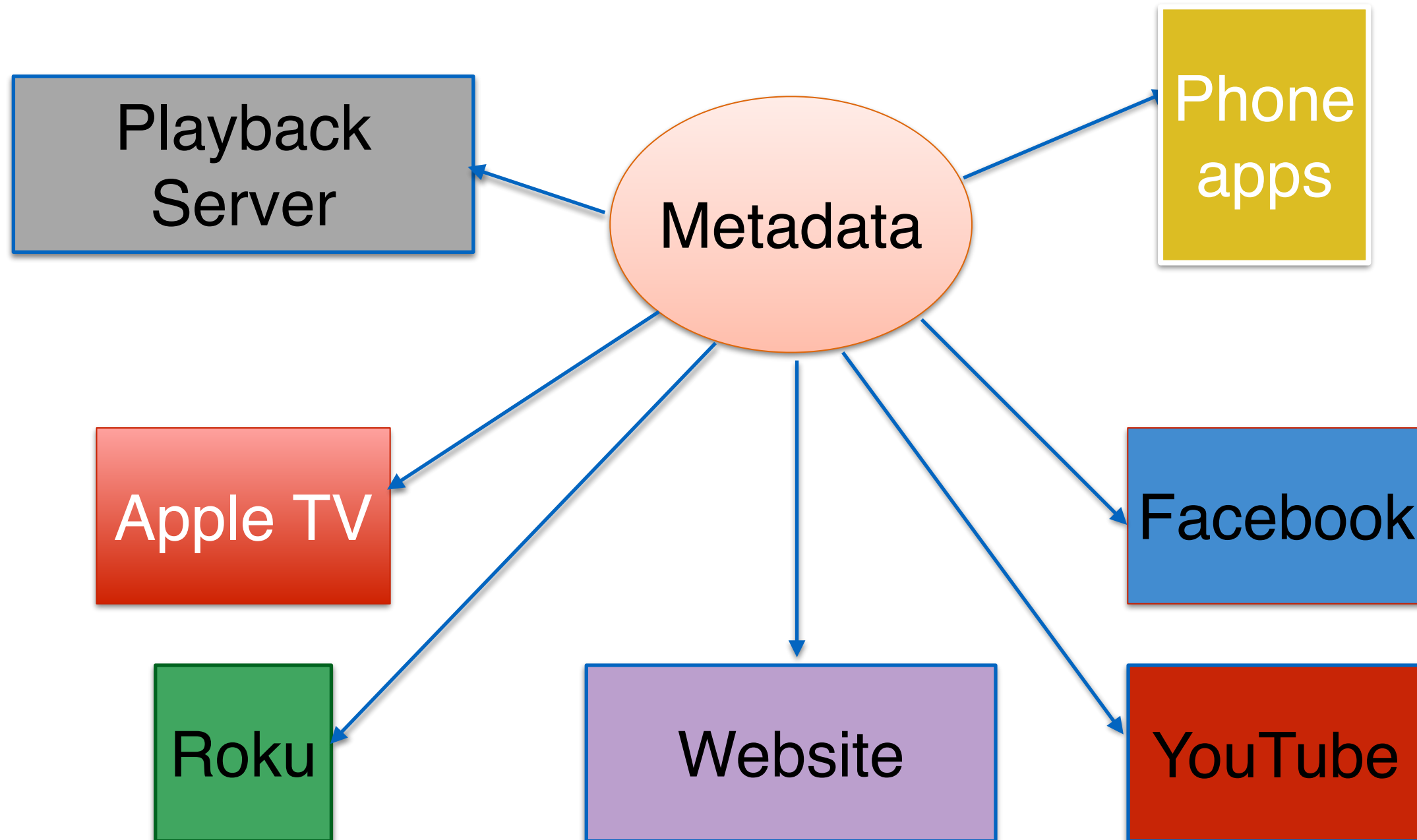
Page Title * CEB 050219 ✓	Drag and Drop More Files Here or <input type="button" value="+ Select files to add"/> <table border="1"><thead><tr><th>Name</th><th>Size</th><th>x</th></tr></thead><tbody><tr><td> CEB050219.mp4</td><td>1.9 GB</td><td></td></tr></tbody></table>	Name	Size	x	 CEB050219.mp4	1.9 GB	
Name		Size	x				
 CEB050219.mp4		1.9 GB					
Page URL * https://archive.org/details/ceb050219 ✓							
Description * ✓ City of Pasco Code Enforcement Board Meeting for May 02, 2019							
Subject Tags * ✓ Washington, City of Pasco, Pasco, Government Meeting, Code Enforcement Board, Viebit, 2019							
Creator City of Pasco ✓							
Date 2019-05-02 ✓							
Collection * Community Media Archive ✓							
Test Item No ✓							
Language Language of the work ✓							
License No license selected ✓							
More Options Add additional metadata...							

Upload and Create Your Item

Spreadsheet Interface

Item	Title	Description	Runtime	Subject1	Subject2
Video1					
Video2					
Video3					
Video4					

Metadata “Hub & Spoke” Model



Specific Metadata Tips

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

EA — Expand Acronyms!

- You now have a Global audience as well as the local audience!
 - Add the long form of the acronym as well
- BOMA - Board of Mayor and Aldermen
- 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

“Pretty Good” Metadata Practices

- Include a “**producer**” element
- include a “**series**” metadata element
- include a “**runtime**” in HH:MM:SS format
- use multiple “**subject**” elements
- put year in a separate “**subject**” element
- put station name, initials and state in separate “**subject**” elements

More “Pretty Good” Practices

- upload **only locally produced** video
- use a **prefix** on your item **identifiers**
- include as much metadata as you have
- learn archive.org’s admin interfaces
- download/backup your metadata - **vendor neutral**

Metadata Recommendations

- Consider the **viewer's needs**
- Consider metadata from the **collection** perspective vs. the **item** level
- Include data only you would know and probably have - **producer**
- Use Consistent Naming in **Titles** (spelling & CAPITALIZATION!)
- Add Dates or Episode numbering to Title - make it **unique!**

- Use a **vendor neutral** solution

- Tips on “Working with the Internet Archive” <https://cutt.ly/mafGte>
- My Metadata “Cheat Sheet”: <https://cutt.ly/0ak7n2>
- These slides: <https://cutt.ly/BtgkqoA>
- Community Media Archive wiki: <https://cutt.ly/XsyOil>
- [Documentary video](#) from 2013 of the Internet Archive
- My email: jhauser@pobox.com

