

# Hosting & Sharing your Media on Internet Archive

John Hauser, Access Humboldt

Sept 3, 2015

Tightrope Media Systems Webinar Series

<http://goo.gl/WcWVXu>

# Scope of the Internet Archive

- nonprofit digital public library - 1996
- 16+ million (books, videos, audio, live music)
- 480+ billion web pages in the Wayback Machine
- Goal is “Universal Access to All Knowledge”
- Documentary about the Internet Archive

# Community Media Archive

- a collection hosted by the Internet Archive
- setup about 6 years ago - attempt to solve VOD for Access Humboldt
- 45,000+ videos, 41,000+ hours
- 44 Access Centers have contributed
- 4.6 million downloads
- 85TB of “original” video files (not “derivative” formats)

# Community Media Archive Vision

- A collection of broadcast quality, locally produced shows with sufficient descriptive information that are freely shareable
  - Good!
- Archive as the hub of a sharing/distribution system for broadcast quality video between access centers
  - DoubleGood!

# How do I get started?

- Register an email address and assign a password
  - <https://archive.org/account/login.php>

# How do I get a collection?

- email Collections group ([collections-service@archive.org](mailto:collections-service@archive.org))
  - title of your collection
  - logo and descriptive blurb about your center
  - request it be a sub-collection of the Community Media Archive
  - provide email address(es) registered with the archive

# How do I upload my videos?

- Use the archive's interactive upload interface
  - <https://archive.org/upload/>
  - Double check that your collection name is selected
    - #I mistake!
- Solution: Preset your collection name in URL
  - [https://archive.org/upload?collection=access\\_humboldt](https://archive.org/upload?collection=access_humboldt)

# How do I get a letterboxed MPEG2 derivative created?

- add a collection:broadcast\_ready metadata element
- "More Options Add additional metadata" link at the bottom of the upload form
- name of the field (collection) goes into "key"
- value of the field (broadcast\_ready) goes into the "value" text entry box



# How do I upload HD video?

- video filename must end in “.HD.mov”
- the original HD file will be available for viewing through the archive’s player as well as the 640x380 .mp4 derivative
- the user must click on the black “HD” logo on archive’s jwplayer toolbar to view the HD version. When viewing HD version, the HD logo will turn white
- Example: <https://archive.org/details/FutCommMediaNYSHDYT>

# Metadata and the Archive

- Minimal required – collection, identifier, title, description, subject
- Anything accepted (and retrievable through `_meta.xml` file on the item's “detail” page)
- subject(s) – now called topic(s); what shows up under “Browse by subject/keyword” link for collection
- Identifier must be unique across all 16+ million items!

# Metadata and Archive - More

- File/Identifier naming restrictions
  - A-zA-Z0-9.\_ - no spaces, parenthesis, braces, pound signs, colons allowed in identifiers or file names
- Playback server is likely more permissive

# Pretty Good Metadata Practices

- push all your metadata to the archive
- 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

# How can I subscribe to a collection or series?

- collections and advanced search results have an RSS feed
- every CMA item gets a torrent file
- some Bittorrent clients can take an RSS feed of torrents
- Automagically downloaded when new items are downloaded
- Except when the IA overwrites this capability!

# Alternatives to RSS Feeds containing Bittorrent files

- Alternative 1
  - script to download the torrent files to a “watch” directory
  - Bittorrent client picks up torrents from the “watch” directory
- Alternative 2
  - script to download the MPEG2 and metadata files to a destination directory

# Series found on the Internet Archive

- Big History Project
- Media Edge
- PopTech
- Peak Moment
- Rompeviento TV
- TedTalks
- The Royal Society for the Arts
- Veterans Helping Veterans

# How can I help?

- Improve the metadata of your items
- How will someone find this item? (in a collection of 45k items)
- Take the time to learn Internet Archive's interfaces
  - “Advanced Search”, “Edit”, “History”
  - “Item Manager”, [ias3upload.pl](https://ias3upload.pl) (bulk uploader)
- Help underwrite work on the distribution aspect of the CMA



# 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 Tightrope Media Systems for hosting this PEG Experts webinar
- To the Digital Bicycle effort from ~ 10 years ago!
- Let's aim for "DoubleGood!" instead of settling for "Good!"

# More Information

- These slides
  - <http://goo.gl/WcVVXu>
- Community Media Archive wiki
  - <http://goo.gl/Wx5m8>
- [john@accesshumboldt.net](mailto:john@accesshumboldt.net)