

Internet Archive Metadata Cheat Sheet

ACM Workshop Version

Title:

make the title of your video unique - include series, episode or date info in the title

Reason 1: to identify this instance of the same titled video

Reason 2: Youtube -> Archive project uses title as identifier; if it sees the same title, bad things are going to happen; only one details page is going to be created but which video is going to get uploaded?

video and metadata may get out of whack, multiple versions of the same metadata elements may appear

Description:

put episode information in the description field before generic series boilerplate.

Reason: search engines index only the first 180 characters of description. This was true as of early 2016.

Metadata for the video should answer the basic Who?, What?, Where?, When? questions. What's unique/specific about *this* video instead of *this other* one?

Potential Metadata fields:

State, City, Center Name (both official "long name" and "short name" or Call Letters), Category, Genre, Series information for each video. These should be entered as **Subject** fields or tags.

You are trying to place the video in "geographic space" as well as time, and subject categories

People's Media Collection specific fields suggestions: "Free Library of Philadelphia", Philly Neighborhood or Branch Library, Class or Workshop Session# and Year, Knight Foundation Program Grant Year info.

Reason: Grouping videos is possible/easier when all available info is at the item level, not embedded in a website's user interface elements

Video can be removed from the website/repurposed without losing the grouping criteria

If metadata uses acronyms that locals would recognize, include the long format version as well for the benefit of non-local searchers: KCB = Keep Cincinnati Beautiful, GTEI = Gateway to Economic Inclusion, CCA = Citizen Complaint Authority

Avoid Reusing Acronyms: CAC=Cable Advisory Commission but in the same collection, CAC also = Citizen's Advisory Council (School Board)

Creator/Producer:

Ensures credit is given to the individual producer of the video even though the Access Center collection aggregates the videos. Enables search of all of a producer's videos on Archive.org

Language:

"English" unless the audio is some other language. Look up the ISO 639 language code for other languages. If metadata uses accented characters, include the ASCII version of the term as well; **archive.org search is not multilingual**

Internet Archive Metadata Cheat Sheet

ACM Workshop Version

Date:

The "Date Published" field on archive.org will be the date the item was uploaded to Youtube or Vimeo

Year will be extracted from the Date field

This generally works OK except for cases where a backlog of old videos from prior years was uploaded at once. In that case we look for a "year" date fragment in the description or title fields and use that to populate the year field

The Year field is important because subcollections of the Community Media Archive are "faceted" by "Year"

Runtime:

Nice to have because users can see the runtime of the video without having to click on the player and the video start to play before finding out how long a video is.

If you had 30 minutes to watch, wouldn't you like to be able to know that the video's runtime is over 2 hours long?

Expressed as HH:MM:SS, example Runtime 00:27:49

Licenseurl:

Consider adding Creative Commons licensing information to your video.

Misc. Notes:

Have a process for indicating a media maker's video should not be included in web distribution

Have a process for removing an item upon request

Archive.org **requires** 5 fields: **Title, URL, Description, Collection, Subject (more than 1 subject allowed and encouraged)**

Archive.org **accepts** any metadata fields you have or want to store. Anything you put in (and edit), you can get back out in JSON or XML format.

To create an MPEG2 version of your video, add a **collection** tag with the value **broadcast_ready** The Archive will automatically derive an NTSC Standard Definition 4:3 aspect ratio video that can be ingested into a playback server.

Links:

Community Media Archive Wiki

https://accesshumboldt.net/wiki/index.php?title=Community_media_archive

Community Media Archive

https://archive.org/details/community_media