

Metadata for Searching Moving Image Collections Comparison

Paul Dougherty 3/7/2016

Metadata for Moving Image Collections, NYU MIAP

The two collections compared are the Pacific Film Archive and the Internet Archive moving images collections (part of Archive.org). The Internet Archive is very large because of its open-ended, crowd-sourced nature. It is also a non-profit entity with a vast volunteer aspect akin to Wikipedia. All this translates to a rich but less than streamlined user experience. The richness is good for savvy computer (science) type users but ordinary users might find the interface wanting compared to those of other institutions and commercial sites. For all these reasons plus the centrality of this collection, extra attention will be devoted to Archive.org (shortened henceforth to IA).

1) How easy was it to find what you were looking for? Was there good documentation about how to search?

Only the first of the four searches will be described at the beginning of this assignment. It will provide a gateway to addressing the first and other questions. The remaining three searches will be discussed the end.

The first search topic is a person, Wallace Berman. He was the subject of a 2007 exhibit at New York University's Grey Gallery. The exhibit page describes Berman as the "The quintessential visual artist of the Beat era."

The Pacific Film Archive (PFA) URL provided by the assignment specs has a very straightforward search box next to a Submit button. This seems to work in a satisfactory way for most general searches. There is text below it with a somewhat faint link to the Advanced Search option.

"For advanced search options, visit oskicat.berkeley.edu. To limit your search to the holdings of the PFA Library and Film Study Center, select "Pacific Film Archive" from the "Entire Collection" pull-down menu."

This link goes to a Quick Search page for OskiCat UC Berkeley Library Web Catalog. The PFA is a subset of this catalog. There is no menu choice for the "Pacific Film Archive" but perusing the long list of options will reveal the (correct) option "Berkeley Art Museum and Pacific Film Archive." At the top of this Quick Search page are several links to the OskiCat Advanced Search page. This page is a robust, full featured Advanced Search page. Up top there is a Help link to a very detailed Search explanation page. It has eight sections with names like "Modifying Searches" and "Preferred Searches."

The Quick Search option on PFA delivered a moving image result entitled "American avant-garde film, 1947-1986." This is from Treasures IV "a two-part set of avante garde (sic) films representing four key decades of American avant-garde filmmaking." This compilation contains the one film that Wallace Berman made called "Aleph" (1966).

At the URL provided in the assignment for the IA "Moving Image Archive" there is a small search box that says (inside it) "Search this Collection." The location of this all-important feature is off to the side and probably not as prominent as it could be. A search for Wallace Berman delivers the Aleph (short) film as well as a charming (edited) home movie entitled "Street Fair 1959" by Edward Silverstone Taylor. It is described as "1959 street fair, upper Grant Avenue, San Francisco--the center of Beat culture." The description lists notable people who appear in the film, among them "artist Wallace Berman."

The Advanced Search on IA is especially robust as it also allows searching on a very long list of metadata codes like `copyrightexpirydate`, `stripped_tags` and `scanningcentre`.

2) Were there multiple ways of searching (e.g. simple and advanced)?

As per the above, yes.

3) Was the search interface easy to use? Was both browse and search supported?

On IA the Advance Search feature was hard to locate. From the Moving Image Archive page (provided in the assignment) there is only the simple "Search this Collection" box. Only on the main Archive.org page is there a link to Advanced Search, this seems like an important omission. A smaller critique is that if one is using the Advanced Search page to find moving image content, the term for media type is "movies" and this might be misleading and mistakenly read as feature-length cinema. In fact it means all moving image materials.

The browse mode of IA seems satisfactory, allowing for the vastness and variety of the collection. But this variety led to certain browse anomalies and it's not clear if they are avoidable. For example the main "buckets" or umbrella categories are listed as image thumbnails akin to a slide table with names like "Television Archive," "Spirituality & Religion" and "Community Video." Going to the Television Archive page will list the collection or contributors on the right, such as BBC and CNN. Whereas Community Video lists topics (on the right) that are sorted by number of such videos that align to that topic within that collection (the highest numbers at the top). The Spirituality & Religion collection lists the topics alphabetically on the right.

It's hard to say how useful it would be to browse the entire PFA collection alphabetically or by date but it is not clear how one would access that. However using one search result (related to Wallace Berman) delivers a record with the following Subject and Genre/Form links which allow one to browse on the basis of various themes.

SubjectFairs -- California -- San Francisco.

Genre/Form Personal/independent works.

Home movies.

Unedited footage.

At the top of the page one "workaround" for browsing the entire collection (say starting at the beginning) is to search for titles that beg with "aa" or "ab."

4) Did you get the results you were expecting? Did you get what you were looking for?

It was delightful to discover the "Street Fair" film. A MIAP graduate posted it on IA and provided the metadata that allowed its discovery as connected to Berman. But that story is for another time. It did seem strange that "Big Sur" delivered no results at PFA.

5) Could you tell what metadata standard was used and if so which one? Is there documentation about the metadata? Make sure you look around the site to see if there is any documentation. If it doesn't say see if you can determine which.

According to the Director of Web Services at the Internet Archive (Rossi, 2011) "IA's metadata schema is based on Dublin Core, but it is extremely flexible." The PFA uses METS and is listed on the "METS Implementation Registry."

6) How well did you think the metadata supported finding what you were looking for?

This is discussed above and in the Addendum.

7) After doing a search were the search results displayed in a meaningful way? Could you tell how they were ordered and was it easy to sort through them?

The PFA results allow one to sort by relevance, date and title. Whereas the IA allows results to sort by Relevance, Views, Title, Date, Archived and Creator.

Addendum (for remaining searches).

Search IA for "Big Sur" (in quotes) delivered 31 results. They are most meaningfully displayed as icons. About two thirds would seem to offer an obvious connection to the topic. Though at a glance it is hard to determine what collections they are from, but by "mousing over" the title of each will cause a "collection" label (such as Community Video) to appear over the thumbnail. When shown as a text list, one only sees the person or entity that posted the video, which is less useful information. Over on the PFA "Big Sur" delivered no moving image results.

Esalen Institute was the next topic, it's the "self actualization" or "human potential" center started near Big Sur in 1962 . It delivered one result at PFA.

"West of center : art and the counterculture experiment in America, 1965-1977 / Elissa Aulter and Adam Lerner, editors ; foreword by Lucy R. Lippard."

The IA delivered seven results and it is pretty clear to see that three are on target while the other four are very peripheral to that topic.

Next a broader topic "computer animation." Using the Advanced Search at PFA for "computer animation" resulted in 111 results searching on Any Field. Often the result would be any clip (say a television show or a education program) that *used* computer animation. Searching for computer animation as a Subject, then the results would be all about or specific to the topic per se. Searching for "computer animation" on the IA turned up 185 results, most were accurate but several animations were vintage. One has to suspect sloppy metadata created by volunteer video posters (aka crowd-sourced video).

References

- Grey Gallery (2007). Semina Culture: Wallace Berman & His Circle. Retrieved from ` <http://web.archive.org/web/20160107074640/https://www.nyu.edu/greyart/exhibits/semina/semina.html>
- Habing, T. (2015). METS Implementation Registry. *The Library of Congress Standards*. Retrieved from <http://www.loc.gov/standards/mets/mets-registry.html>
- Rossi, A. (2011). How Archive.org items are structured. *Internet Archive Blogs*. Retrieved from <https://blog.archive.org/2011/03/31/how-archive-org-items-are-structured/>