

Media Finders: Expert Search Intermediaries for the Online Catalog
An OLAC Conference Poster Presentation
Kelley McGrath
Ball State University
kmcgrath@bsu.edu

What is a Media Finder?

A Web form that serves as a dynamic pathfinder for searching some subset of materials in the OPAC.

- Uses cataloging knowledge to present the user with options that are used to construct searches in the OPAC
- Provides guided access to materials that have characteristics not adequately served by a generic search box in a standard OPAC

Why make a Media Finder?

Browsing access

- For unclassified, closed stacks, which is our situation for media at Ball State. If the user can't find something in the catalog with a keyword search, he or she is out of luck. In our catalog (Sirsi) searches cannot be limited by format in browse mode in the public interface.
- For remote users
- For browsing by multiple aspects of an item, e.g. a 1956 foreign comedy can be browsed as part of the category of foreign comedies released during the 1950s

“Blank screen” problem

- Many users don't know exactly what they want or how to start looking. Media Finders give an idea of what the library has and what sorts of characteristics can be searched.

Fuller use of information in the catalog record

- There is lots of information in the bibliographic record that could be useful for searching but most users do not have either the search skills or the inside knowledge necessary to construct the search strings

Automates long and/or complex search strings

Improves access to targeted subsets of materials

- Increases precision by using background limiters in all searches
- Provides an easy way for users to access information “hidden” in bibliographic records without having to know how to construct complex searches

How to make a Media Finder:

A Media Finder consists of three elements:

1. Standard HTML Web form

- Presents the user with a set of predefined options and outputs the user's choices to the CGI script
- Easy to design; uses standard input methods such as pull-down menus, checkboxes, radio buttons, and text boxes

2. CGI script
 - Simple program that converts the user's choices into a dynamic URL
 - HTML form only outputs bits of data such as genre="western" and captioned="yes"
 - CGI script takes these bits of information and puts them together to form a dynamic URL
 - Can be written in a variety of languages such as PERL or ASP
 - Mainly consists of basic text processing and if-then functions so the CGI script is fairly easy to create in whatever programming language is used

3. Dynamic URL that searches the OPAC
 - Example:
[http://liblink.bsu.edu/uhtbin/cgisirsi/x/SIRSI/0/5?searchdata1=\(feature+OR+short\)+films{su}&item_2cat=VIDEO](http://liblink.bsu.edu/uhtbin/cgisirsi/x/SIRSI/0/5?searchdata1=(feature+OR+short)+films{su}&item_2cat=VIDEO)
 - The part before the question mark is static and in our case indicates the location of our OPAC's server and information about the type of search (keyword)
 - The part after the question mark is dynamic and communicates the search terms

In order to set up the search options and construct the URL, it is necessary to be familiar with:

- Your collection and local cataloging policies
- Effective search strategies and Boolean searching
- Indexing and search policies in your OPAC (Note that Sirsi has an implicit SAME operator so in our searches, when there is no operator between terms, it is implied that the terms will be in the same MARC field)
- How to construct a URL that will search your specific system

Evolution of the Media Finders at Ball State:

1. Video Finder (fiction and drama)

This was the first Media Finder developed because it was the easiest for me to do. Most of the information I wanted to use for retrieval was already in the record and the collection was big enough to be interesting, but small enough to be manageable (around 2500 records). The Video Finder uses the subject headings "feature films" or "short films" and the format "videorecording" (from the item record) to delineate the desired subset of all our videos (most of our video collection consists of documentaries and other non-fiction).

Search options include

- Free text box
User types in search terms; defaults to general keyword, but author or title keyword also available

- Format (pulldown list)
Based on information in item records in Sirsi. Could also be done based on information in the Type fixed field (leader 06 - Type of record) and 007 if your ILS allows keyword indexing of these fields. (Sirsi does not.)

- Major genre headings (checkboxes)
These were mostly already in the records although we have not always been as proactive in adding these headings as we might have been.

- Captioned and audio-described videos (checkboxes)
Based on the “Video recordings for the hearing impaired” and “Video recordings for people with visual disabilities” subject headings which have consistently been added to records when appropriate.
- Awards (pulldown list)
Keyword search on the 586 field.
- Country of origin (pulldown list)
We retrospectively added geographic subdivisions for country of origin to the subject headings feature or short films for non-U.S. films where we were able to identify them by looking for non-English languages in 546 (before we had a significant DVD collection) and by reviewing records in the catalog. We also ended up adding the subject heading “foreign films” to these films and adding a foreign film checkbox to the list of genres. This was our most requested enhancement.
- Original release date (pulldown list)
This is a good example of information that is important to users and typically not easily accessible from the basic catalog interface. This search is based on a standard note about the original release or broadcast date. We currently use the phrase “motion picture” in most notes, but I did have to clean up a few old records that used the word “film.” We also had a number of records where the note had been omitted, often apparently because the year of release and the year of publication of the video were the same (which is probably fine for conveying the information to someone already looking at the record, but is not helpful for retrieval).
- Setting—place (pulldown list)
Based on the 043 field which allows access at both the continent and country level. For a long time, we have added 043 consistently when there is a geographic subject heading or subdivision in the record, but are now trying to be more proactive in adding this information for feature films.
- Setting—time (pulldown list)
I would like to expand this option, but without coming up with some local way to add information to the bibliographic records, I'm not sure how to do this. It's hard to do something consistent with the current information in our records since historical periods in subject headings vary between countries and often, even when a historical period is expressed in the 520 summary, no subject heading has been added. Currently, the options I have provided are mostly wars or events rather than time periods.

Probably I could do something like “nineteenth ADJ century {su} OR 18??{651}” at the century level, but there would be some false hits (if the historical period is 1895-1925 [made up example] and the film is clearly set in the 1910s for example) and a lot more things that ought to be picked up, but won't be because no subject heading has been added.

2. Non-Classical Music Finder

This search page and the sounds search page were developed in collaboration with Sue Weiland, Ball State's music cataloger. The options on this search page are focused on non-classical genres although we don't actually have any way to eliminate classical music from the results. Unlike the Video Finder, these searches are mostly based on subject headings. We provided the ability to limit by decade under the genres where this is allowed for in LCSH. In addition to genres we also included searches for world

music by country (again based on subject headings) and topic (such as drinking songs or Christmas music).

Two new features were added to this search page:

- Ability to choose whether to use the Boolean operator OR or AND when searching multiple genres. This was triggered by a discussion of two situations in which opposing options are best:

Rap AND contemporary Christian Spirituals AND gospel	vs vs.	Rap OR contemporary Christian Spirituals OR gospel
---	-----------	---

Since there is no single best option, we decided to let users pick and hope that they're not too confused. The default is AND.

- Ability to choose different options for sorting results, i.e. by title, author or publication date. With videos, title sorting was the only thing that made sense (date of publication probably being more misleading than helpful).

3. Sounds and Sound Effects Finder

This search page provides access to a collection of production music and sound effects used mainly by telecommunications and theater students. It also includes animal sounds and other nature sounds. Like the non-classical music finder, this page is mainly based on subject headings. It's a very basic search page, but does give some prominence to a set of materials many users probably don't know we have.

Future possibilities

Classical music

Sue Weiland, Ball State's music cataloger, and I have tried to develop a Media Finder for classical music, but our attempts have run up against a couple of significant obstacles that we have not been able to overcome.

1. Multiple works with multiple characteristics on one bibliographic record

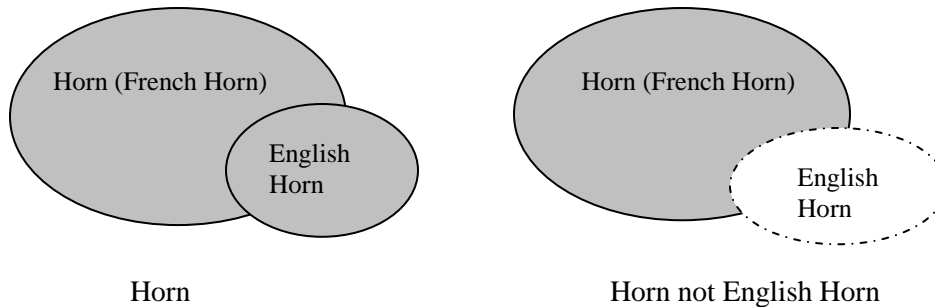
Many classical recordings contain numerous pieces in various genres, with various instrumentation, by various composers. This information is not connected i.e. there is no link between subject headings (genre and instrumentation) and uniform titles (composer's name and the title of the piece) (at least for non-generic uniform titles). Therefore, if the user wants Beethoven symphonies, there is no way to exclude recordings with, for example, a Beethoven sonata and a Mozart symphony.

The general problem of multiple works on one bibliographic record with no connections between related pieces of information about a given work can potentially affect some of our other search pages, but the problem is so far most pronounced with classical music. While bizarre results do occasionally occur with films on video (especially the really old ones) and pop music recordings, they are less prevalent (most videos contain only one film and most pop recordings are either by one artist or in one genre) so the problem is not as noticeable.

Perhaps the implementation of FRBR will be able to help with the problem of searching multiple works on one bibliographic record.

2. Undifferentiated keyword terms

For example, “horn” is used for French horn, so there is no keyword search that will get all and only French horn pieces (“Horn” gets French horn and English horn and whatever other horns while “horn not English horn” gets recordings with French horns and no English horns but excludes recordings that include both French horns and English horns”)



Another example would be the LC subject headings “Songs” (used for secular songs) and “Sacred songs.” From a keyword searching point of view, there is something to be said for all headings having some sort of qualifier to differentiate them.

Audiobooks and spoken recordings

We are fairly close to having something functional based on a combination of information from the Ltxt fixed field (field 006, 13-14 - Literary text for sound recordings) (which we had to transfer to a field that is searchable in Sirsi) and genre headings.

Biography

I have tried to use searches that combine the subject subdivision biography with subject headings for groups of people in a broad category (such as music or literature), but since the groups of people tended to consist of a rather long “OR” searches, I ran up against Sirsi’s 250-character limit on keyword searches. This is not likely to be a problem for users typing in searches, but it does cause problems with artificial searches. I am hoping to take the result sets based on the subject heading searches and add some sort of category code to the records to decrease the length of the search.

Fiction and literature

I would like to do something using the LitF fixed field (field 006, 16 - Literary form) and genre headings. It is hard to construct a background search because a lot of our older retrocon records have 0 in the fixed field when they should have 1. Also, most of our drama and poetry is not coded in the fixed field as this option only became available in MARC 21 recently. Fixed field coding is also limited to one form of literature per record and is not helpful with multi-form items.

Non-fiction videos, pictures, posters, 3-D, software, and other media

We have a lot of this stuff (also unclassified and mostly in closed stacks) which would benefit from better access and especially some sort of broad subject categories to aid browsability. Keyword works reasonably well for a focused search, but to see all the art videos, for example, a user would probably have to make a long OR search and wade through a lot of false drops.

Archives and special collections materials

I have a draft page (<http://web.bsu.edu/kmcgrath/mf/spcollmedia3search.html>) that browses or searches within specific collections of materials and allows limiting to local history or by format (video, audio,

etc.). Our archivist is concerned that if users see this page, they will think they can find everything from there rather than just the subset of materials available through the catalog so we have not yet posted this page.

Other Miscellaneous Thoughts

- Descriptive notes vs. information for retrieval

A lot of information is in bibliographic records for description and recognition purposes for users looking at the record that could, if more consistently coded, be used for retrieval. This is especially true for audiovisual materials where a lot of information is only in note fields.

- Faceted browsing

A great example of the possibilities of iterative faceted browsing is the Flamenco Project at Berkeley (<http://bailando.sims.berkeley.edu/flamenco.html>). In this interface, users are able to see their options up front and they are prevented from retrieving zero result sets. I think the future of effective online browsing lies in making fuller use of facets and would like someday to develop something like this to browse our collections. With the Media Finders, users have to guess at what combinations of options actually lead to results and there are plenty of combinations that produce zero results.

More Information

Ball State's Media Finders are available at: <http://www.bsu.edu/library/librarycatalogs/mediafinders/>

The Video Finder was first posted in September 2003 and the Non-Classical Music and Sounds and Sound Effects finders were posted in December 2003. Since December 2003, we have averaged 606 hits per month on the video page, 217 hits per month on the music page, and 55 hits per month on the sounds page.

Email me at kmcgrath@bsu.edu if you would like a sample of the ASP code behind the scenes. Because the code is adapted to work with our implementation of Sirsi and with our local cataloging practices, it will need some adjustment to work in other circumstances, but the general principles can be applied to many situations and materials.