Open Folklore: project and website review

Perry Willett, California Digital Library

January 7, 2014

I want to thank Julie Bobay and the others associated with Open Folklore for the opportunity to review the OF Project. Unfortunately, I was unable to attend the meeting in Bloomington on July 8-9, 2013 and missed out on the discussion. I didn't feel like I should comment on the report by the main body of reviewers[1] having missed the meeting and discussion. Julie invited me to contribute a separate report, which I happily supply.

I appreciate the experimental nature of much of this, and am impressed by the efforts of the AFS and IU Library staff. There are no obvious answers about which directions to take or services to provide. As a general rule, any website needs to have a clear idea about its audience and the information and activities provided for them. There has to be a reason for someone to go there, and then to return.

I'm basing the following analysis on my own interactions with websites, and my own sometimes painful experience in building websites. I've learned the hard way too many times that people don't come back to a website out of general interest or curiosity. They may come the first time, but they'll decide pretty quickly if it's worth coming back. I know how hard it is to hear criticisms of work like this, and how hard it is then to decide whether the critiques are correct. I offer the following analysis in the spirit of improving the OF website.

On the whole, the website offers a good outline about open access and folklore, and I like the design very much. It's streamlined and inviting. However, there's very little for anyone looking for more information. The project briefing we received contains a lot more information about what's going on behind the scenes. I encourage OF to include more information to the website without overloading it or larding up the clean look—perhaps linking to secondary pages with more information.

The announcement about the new version of OF stated that:

> Since 2010, Open Folklore has operated as a) a website; b) a scholarly portal providing access to open-access books, journals, websites, and gray literature in our field; c) a branding effort or unifying label for a collection of projects and efforts being pursued by AFS and IUBL to make a greater range of scholarly resources in folklore studies openly available for those who need them; and d) a case study for productive and effective collaboration between an scholarly society and an academic library.

Let's look at each of these.

---

[1] Posner, Miriam; Sims, Nancy; Watrall, Ethan. Open Folklore: Maintaining Momentum, Assuring a Future. American Folklore Society. July, 2013. Accessed 2014-Jan-07. http://hdl.handle.net/2022/17192

A. Website

As a website, OF provides links to news, blogposts and tweets about open folklore. However, these are pretty sparse. There is only one news item from all of 2013. If this is an important aspect of OF, then the news feed and blogposts need more frequent updates.

It's damaging to have a sporadic and out-of-date newsfeed and blog because it sends a message that the site is not maintained. People will notice this quickly and won't come back. This is not a trivial amount of work, so careful consideration is needed about whether the OF team has the capacity. I'd recommend removing this news and twitter feed if the resources aren't available to update them regularly. If you're looking for subject matter for blogposts, I note that a number of the projects described in the project briefing would make very interesting blogposts (e.g., the situation with scanning the IU Folklore collection and contracts with Google).

I'm not clear on how OF relates to the AFS website. The AFS newsfeed and calendar are much more active and current.

The AFS Ethnographic Thesaurus is included as a separate page, but without a clear purpose. It's not tied to search of anything—one can find subject terms in the thesaurus, but you can't do anything with them. Is it available for download? Is there an API? If you want people to use it by incorporating it into other services, you have to provide a way for them to either obtain it or link to it programmatically as linked data; MeSH may provide a model for this:

- http://www.nlm.nih.gov/mesh/filelist.html

I ran the website through an automated test for accessibility, to test how it would do with assistive technologies such as screen readers. On the whole the website came out well. The main failure was a lack of "alt" attributes in some <img> elements. This is a significant problem for people using screen readers and should be addressed. See the spreadsheet for more details (OFAccessibilityReport20140106.xlsx). The SSB Bart Group website is a good source for information about accessibility: http://www.ssbbartgroup.com

B. Scholarly portal

The portal pages on books, journals, websites and gray literature provide information about each of these formats. Each of the pages contains more information with links to important resources, along with a search feature. Let's look at each of these:

1. Search

Most people visiting the OF website will interact with it via the search feature. However, it's not at all clear what is being searched, and there is no explanation available on the site. Here are some of the problems I discovered:

- Haphazard coverage: unable to find articles in journals listed in the OF journal page: <http://openfolklore.org/journal-list>. And the same for books—the books page specifically mentions both "Grimm's Household Tales" and "Motif-Index of Folk Literature" but they can't be found via the search box. Much is made (appropriately) of Unglue.it's open access

publication of "Oral Literature in Africa" by Ruth Finnegan. However, I can't find it with Search either. Any title that's used as an example in the website should be available via search.

- Incorrect metadata: A book listed on the book page, "The Anguish of Snails," is listed as an article in the metadata. More significantly, the metadata for 2 books by Stith Thompson, "Round the levee" and "The folktale" list their publication date as 2013, although they're from the 1930s and 40s.
- Multiple metadata records: The two books mentioned above, "Round the levee" and "The folktale," are listed multiple times. (The former does have two copies in Hathitrust, but is listed 4 times.) I saw a lot of titles with this problem.
- Odd behavior: after typing something into the search box and clicking on "Search," alternative search terms appear. It's odd for this to happen when clicking on "search." One has to click on search again to conduct the search. Maybe this is a feature of Drupal. Not major, just a little annoying.

Recommendation:
- At the very least, add information about what is being searched. It would be better to improve search, since this is how most visitors to the site will interact with the content. It's important that it work well. As it stands now, it's just not good enough.

2. Books

The page contains two statements of when to use OF for books, but neither one is very helpful. The two "when to use OF" lists push other information "below the fold."

Given the state of copyright, it's difficult to know whether any particular book may be "open" until someone looks for it. This puts the researcher in a difficult position. He or she is looking for a book, but doesn't find it in OF. Is it not there because the book is out of scope, not open access or public domain, or because of problems with search (see above)? This is a very hard problem. I would recommend removing both of the "when" lists and replace it with a "what's here" list, to complement the (very useful) "what's not here" list.

Some other observations on this page:

- Why is there only the "partial folklore collection" in HathiTrust? There's actually an interesting discussion of this in the project briefing. This would make an interesting blog post.
- There are multiple mentions of collections and projects on the page, but they're not always linked. Add links:
  - To the open access books from Utah State Univ Press (at the top of the page).
  - To the IU Folklore collection in HathiTrust (at the bottom of the page).

Recommendations:
- Remove both of the "when" lists and replace it with a "what's here" list.
- Complete the video tutorial (with captions) on using the Open Folklore website to discover open access books.

3. Journals

This page is very informative and useful.

Recommendations:
- Information is given for journal editors, but most researchers are authors of articles, not journal editors. Add information about open access for authors of journal articles.
- The video, while very useful, should provide a captioned version for people with hearing impairments.

4. Web archives

Looking at the collection of websites, I would make a few observations:

- It seems experimental (which it undoubtedly and understandably is). You might be more open about your plans for this part of the collection.
- Most of the websites have only been harvested once. For some sites, such as the American Memory "Buckaroos in Paradise" collection, this makes sense and works well. For more active sites, such as the AFS site, crawls need to be done more frequently. It makes no sense to have one or two crawls from 2010 of actively changing websites.
- For the American Memory sites, none of the archival content (photos, etc) was captured. So what you're left with are the essays and other contextual information provided by the site. Since the Library of Congress is committed to the American Memory content and site, it's pretty safe compared to other websites. The American Memory collections seem like a lower priority for preservation in Archive-It.

Web archiving is also not a trivial undertaking. It takes a good deal of work, particularly post-harvest, to determine if the site was captured adequately. You need to be sure css, javascript, image, and other auxiliary files were captured, and if not, determine why.

Recommendations:
- Create a collection development plan for the collection, to determine types of sites to include and frequency of capture.
- Add a feature to allow people to nominate websites to be added the collection. See the Nomination Tool at the University of North Texas Library: http://digital2.library.unt.edu/nomination/
- Provide more explanation about plans for this feature.

5. Gray literature

This is an important format for materials in this area. The page and presumably searchable metadata point to only a single repository. There is room for a great deal of expansion here.

C. A collection of projects and efforts pursued by AFS and IUBL

The various projects described in the Project Briefing are exciting, maybe more so than content available through the website analyzed above. These efforts need to be described on the website—perhaps, as I suggest above, as blogposts.

The description of the work with Hathitrust in researching rights status of works and getting complete collections from Google, sounds very exciting, and to my mind one of the most valuable accomplishments of this effort. I encourage you to continue this work.

The advocacy work for open access is extremely important. As the meeting notes describe, there is a great need to educate the folklore community about open access, and the seminars, training sessions and other presentations are all extremely valuable.

Recommendation:
- Include more information about the various OF projects on the website.
- Add a calendar for training sessions, presentations, workshops etc. Of course, this would need to be updated regularly. It's easy for this kind of information to go stale.
- Add promotional and other materials about open access to the website, so that people can use them to give presentations and conduct workshops at their home institutions.

D. A case study for collaboration

Are there benchmarks, metrics, schedules? I realize that this is a pilot project and experimental in a lot of ways, but it's important for your audience(s) to know. If, for instance, you know there are features that you'll be significantly improving, you'll give them a reason to come back. Otherwise, they'll assume that nothing much is happening.

Finally, let me end by recommending that you add much more information for folklorists about open access. There is encouragement given for researchers to contact OF about making their books and journals freely available on the "formats" pages. This should be given more prominence, with a separate page about open access. For instance, there's no information given to researchers about making their journal articles open access, but this may be the simplest action they can take. People will have questions about copyright, repositories and open access. There is may be enough information via other professional organizations to simply link to it. You could also point to university sites such as the University of California's on open access:

http://osc.universityofcalifornia.edu/open-access-policy/

You should make your conference presentations and workshop Powerpoint slides available (in IUScholarWorks, Slideshare or some other open repository) and point to them from the OF website. Have blogposts or a calendar for events. You could also create a place for people to share training and publicity materials. There is still a great need for advocacy for and education about open access, and OF could play an important role in the Folklore community in this area. This will require a significant commitment of time, but with a team approach you can divide up the work so that no single individual is entirely burdened.