November 2013

Back Talk -- Cataloging: Integrating Web Resources into OPACs

Anthony (Tony) W. Ferguson
University of Hong Kong, ferguson@hkucc.hku.hk

Follow this and additional works at: http://docs.lib.purdue.edu/atg

Part of the Library and Information Science Commons

Recommended Citation
DOI: https://doi.org/10.7771/2380-176X.3177

This document has been made available through Purdue e-Pubs, a service of the Purdue University Libraries. Please contact epubs@purdue.edu for additional information.
posted to your online forum?" These opening questions on the Chilling Effects Clearinghouse Website seek to help users understand the protections that intellectual property laws and the First Amendment give to online activities. Born of a unique collaboration among law school clinics and the Electronic Frontier Foundation, the site offers and solicits additions to a searchable database of “Cease and Desist” notices as well as information on copyright and fair use; “fan fiction” (authors writing stories featuring characters from other copyrighted works); anonymity, hyperlink protocol, protest, parody and criticism sites; and trademarks. (Edpage, February 25, 2002) — http://www.chillingeffects.org/

Science

Remember “On Top of spaghetti, all covered with cheese” written and recorded by Tom Glazer? A lyric is worth a 1000 inspirations, and at Science Songs, there are more than 100 tracks organized in major categories of songs about Space, Energy & Motion, Experiments, and Nature, many recorded by Tom. Lalala along with the Eolithips (“dawn horse of course”) song, fly with the Stratus and Cumulus lyrics, or sing along with zoom a Little Zoom or Beep, Beep. The original dusty albums were exhumed from the basement of Jef Poskanzer’s parents’ basement and converted into .mp3 files (note they load and play rather slowly with Windows Media Player). The songs may seem corny, but will definitely appeal to K-12 science educators. As an added bonus, the site links to other music resources including Dr. Chordate (“science meets music for education fun”), and to Scoutsongs.com which includes “On Top of Spaghetti” lyrics and music. — http://www.acme.com/jef/science_songs/.

Reference

One of the most widely used undergraduate textbooks in molecular and cell biology, Molecular Biology of the Cell by Bruce Alberts et al., was the first title made available on the National Center for Biotechnology Information Website as part of their Entrez retrieval system for searching linked databases. The databases include: PubMed for biomedical literature; Genbank’s Nucleotide sequence database; Protein sequence database; Structure: three-dimensional macromolecular structures; Genome: complete genome assemblies; PopSet: population study data sets; OMIM: Online Mendelian Inheritance in Man; Taxonomy: organisms in Genbank; Books: online books (full text and chapters); ProbeSet: Gene Expression Omnibus (GEO); and 3D Domains: domains from Entrez Structure. The Books section currently includes 6 full text titles in cell and molecular biology and genetics and one chapter. Searchable by key concept (or one can browse, although this is still a bit clumsy), the contents are also linked to PubMed citations. — http://www.ncbi.nlm.nih.gov/entrez/

I was recently asked while attending the 2002 OCLC RLGAC meeting to comment on the value of sticking with existing cataloging standards—but from the point of view of someone who has focused mainly on collection building and whose collecting horizons go beyond those materials collected by most North American research libraries.

As I thought about this assignment it occurred to me that while libraries have changed significantly over the 30 years I have been a librarian, there have been three interlinking constants which have shaped them and which will continue to insure their value as places for recreation and learning: collecting, organizing, and helping people effectively use them to find the information they need.

I think, however, we are about to experience a paradigm shift in what we collect. Our job as librarians has always been to "buy" as much of the information needed by our readers as our budgets would allow. In some cases this has meant librarians doing the selection, in others it has meant overseeing the activities of selected readers like the faculty who decide what to purchase or not to purchase. But the focus for most of us (excluding archives associated with organizations) has been on exchanging money for content. This is what I believe is going to change. Instead of buying 90 to 95 percent of the content needed and relying upon gifts/exchange for the other 5 to 10 percent (I am being overly generous) for the rest of what we add, I think we are going to must shift to harvesting from the Web the information resources needed by our users and supplementing that with purchased content.

Now I know that at this point many of you will go into “the Web is full of junk” defense mode. Let me suggest that we drop that as self-defeating behavior. Our job has always been “getting” the best content for the use of our readers, not “buying” it. It just so happened that “buying” and “best” usually went together. In some ways the SPARC initiative is another variation on this theme but it is only an attack on the crass commercialization of content. So instead of demeaning the Web to which most of our readers daily pledge allegiance (they vote with their feet and fingers), let’s refocus our energies and organizational passions to figuring out how to identify those Web resources which match the needs of our users. I believe that as we gradually do this, the proportion of purchased resources will gradually get smaller and the proportion of freebies will get larger.

When you look at the history of the great academic libraries of the world, you invariably read about Mr. X or Miss Y who was the first librarian and who was charged with selecting, acquiring, cataloging, and helping users find what they needed. As needs and budgets got larger, more people were added, buildings were built, and complex library organizations were created. In the post-War and post-Sputnik eras, the size and complexity of library structures skyrocketed. What I would like to suggest is, with the advent of the Web, we need to change how we consider a new and expanded universe of publication and stop thinking of Web freebies and roadkill as similar in quality. We need to do this for two good reasons: First, the amount of good material is enormous. Second, it is where our users are at and we can’t afford to be marginalized.

Having said that, let’s turn to the changes that integrating Web resources within our collections will mean to our bibliographic access policies and procedures. A typical defense mode I’ve encountered when discussing putting links from our catalogs to Web resources is quick agreement that librarians should create subject-based library Web pages but a rejection of any thought that we really catalog them like books because they are unstable, we don’t own them, etc. Likewise, suggestions that libraries share the work of cataloging all of these resources are rejected because there are just too many of them and because it would be too costly to use our tried and true MARC 21 and AACR2 standards to catalog things that will disappear or go dormant tomorrow. These reasons, moreover, seem even more reasonable when you consider that we are just now at the birth of the Web. People in the world’s most populous nations have yet to integrate it into their daily lives. Once they do, the Web will really begin to grow. But all of this thinking is beside the point. We cannot do our bibliographic business as usual because our collecting business is not going to be as usual. We have got to get more creative and dare to change.

Allow me to point out that this would not be the first time we did something different/innovative/creative. With the advent of the journal/magazine medium, we (everyone involved in the information chain and not just librarians) figured out a simpler and more cost-effective solution: the index which then evolved to the abstract which has now evolved to full text. Somehow we got over the need to apply all of our cataloging rules and procedures to journals. Indeed, when the number of books got too many for even the super rich libraries to catalog everything, we invented bibliographic utilities like OCLC and RLG. We are capable of great things, we just have to allow ourselves to let loose of the past.

Last fall I was at a meeting of librarians in the Wutai mountains of China where the Dublin Core was explained as a possible alternative to what we are doing now. Of course, that wasn’t the first time I had heard it explained but somehow contrasting it with what we have been doing in the past—to librarians who lacked the resources to follow our lead, it seemed so clear to me then that we needed to adopt a new way of describing collections of content that is simple enough for authors and where appropriate, publishers, to take part so that providing access to the exploding universe of information is possible. We need, I believe, to recognize that the world’s authors of Web content are not going to convert to MARC 21 and AACR2. Unless we want to get left in the digital dust, we have to be willing to work with others to find ways of working smarter not just harder.

This leads me to the need to say a few words about the effect of harvesting Web resources upon helping learners to find the information they need. Several times over the past few years I would ask groups of students in North America to indicate by a show of hands where they went first when they needed information: the library or the Web. The Web always won. Recently I met with a group of architectural students in Hong Kong and the results of my simple survey were the same. My experience, however, has also shown that most students know little about advanced Web searching techniques. They ignore our sophisticated library catalogs, but they are ignorant of how to improve their chances of finding what they want on the Web. We’ve got to refocus our public service efforts to teach learners how to use the Web and then use the opportunity to alert them to print and digital resources that have been purchased to meet their needs.

Lorcan Dempsey, OCLC Vice President for Research, suggested at the RLGAC meeting that we need to “surfacing” library resources wherever our learners are working. I didn’t catch on to what he meant until it occurred to me that what he meant by surfacing was like what happens when you submerge an air...