

11-7-2008

Bringing Semantic Diversity to the Online Catalog with LibraryThing

Rachel A. Erb

University of Nebraska at Omaha

Melissa Cast-Brede

University of Nebraska at Omaha, mcast@unomaha.edu

Follow this and additional works at: <https://digitalcommons.unomaha.edu/crisslibfacproc>

 Part of the [Library and Information Science Commons](#)

Recommended Citation

Erb, Rachel A. and Cast-Brede, Melissa, "Bringing Semantic Diversity to the Online Catalog with LibraryThing" (2008). *Criss Library Faculty Proceedings & Presentations*. 34.

<https://digitalcommons.unomaha.edu/crisslibfacproc/34>

This Presentation is brought to you for free and open access by the Dr. C.C. and Mabel L. Criss Library at DigitalCommons@UNO. It has been accepted for inclusion in Criss Library Faculty Proceedings & Presentations by an authorized administrator of DigitalCommons@UNO. For more information, please contact unodigitalcommons@unomaha.edu.



Bringing Semantic Diversity to the Online Catalog with LibraryThing

Rachel A. Erb
Systems Librarian
University of Nebraska-Omaha

Melissa Cast-Brede
Education Librarian
University of Nebraska-Omaha

Abstract

While controlled vocabularies, such as the Library of Congress Subject Headings, are an essential component of bibliographic classification, a controlled vocabulary excludes all possibilities of semantic variance by design. Also, a controlled vocabulary tends to lag behind the organic nature of language and does not account for the introduction of new or discipline specific vocabularies. These limitations present unique challenges for our users searching the OPAC. Can importing social tags in the online catalog effectively address the lack of semantic variance?

As part of the Web OPAC redesign project at UNO, LibraryThing tags were added to matching bibliographic records in the online catalog. This presentation will cover the practical aspects of adding LibraryThing tags to most vendor-based OPACs, address the variety of tags employed and offer ideas for effective tagging. In addition, we will explore how a collaborative service learning project with discipline specific university classes encouraged patron participation. We will also examine the overall quality and utility of LibraryThing's folksonomy. Lastly, additional features to be added in the near future by LibraryThing's developers will be discussed.

LibraryThing Folksonomy in the Online Catalog

In recent years, there has been experimentation with incorporating social tagging in the online catalog. The rise in popularity of social web services such as del.icio.us and LibraryThing are proof of the public's interest in describing personal collections. OPAC vendors have added this feature to traditional online catalogs, but depending on the system, these features could incur extra costs that are often prohibitively expensive. Moreover, one of the inherent problems with tagging in the OPAC is that it is often difficult to garner enough user participation to generate significant metadata. Because LibraryThing is essentially an online catalog for users to catalog and tag their private book collections, there is great incentive to provide meaningful and relevant folksonomy. LibraryThing.com is a vibrant online community, and therefore, there is plenty of user-generated metadata that enhances existing metadata in the online catalog. It is now possible for libraries to easily incorporate LibraryThing tags in the online catalog for a nominal fee, and therefore connect users to similarly tagged books in their collections.

Folksonomy Tag Studies

Few formalized studies of folksonomy tagging projects in locally customized tagging environments are related to our project and its objectives. The Proof of Concept study conducted at the Metropolitan Museum of Art involved comparing tags assigned by trained and untrained

catalogers. Because approximately 77% of assigned terms from untrained catalogers were descriptive, the researchers asserted “non-specialists can supply a useful number of new access points, augmenting the professional descriptions of art museums” (Trant 102). The successful outcome of this study of folksonomy within the museum catalog context has encouraged the Steve collaboration (<http://www.steve.museum/>) to further investigate and develop social tagging initiatives of art museum collections.

In a study assessing the efficacy of tag cloud searching, students from a first-year engineering class at the Australian National University participated in tagging 10 articles each, entering as many tags as they felt necessary (Sinclair and Cardew-Hall 18). The articles were pre-selected because researchers were concerned participants would assign better quality tags to articles of personal interest (Sinclair and Cardew-Hall 22). The authors found participants assigned relevant keywords which contributed to their overall findings that tag clouds have several “positive attributes” (Sinclair and Cardew-Hall 27).

General characteristics of folksonomy tags have also been analyzed. Most of the research focuses on the tagging environment of del.icio.us; however, these results are relevant in analyzing other folksonomy tags. In a seminal study that encouraged subsequent research, Golder and Huberman’s analysis of 68,668 del.icio.us bookmarks indicated that tags perform several key functions. Tags can: 1) identify the topic, 2) describe what the item is (“blog”, etc.), 3) declare ownership of the item (“author”, etc.), 4) refine or qualify existing categories, 5) identify qualities or characteristics, 6) perform self-reference and task organizing.

In a subsequent study, Munk and Mork analyzed 76,601 keywords from 500 randomly chosen taggers with del.icio.us accounts. They focused on analyzing tags germane to information technology fields because 87 of the most popular del.icio.us keywords are in this field (Munk and Mork 120). They discovered that the distribution of tags follow “the classic power law where very few keywords are dominant” (Munk and Mork 116). They also revealed that taggers with more professional IT expertise, regardless of professional focus, employed professional, IT specific tags. The “casual IT dabbler”, on the other hand, assigned tags “in very broad common cultural categories” (Munk and Mork 116). The general categories of tags include content, media, genre, copyright, value, meta-reflexive (“Mytags”), process (“2read”), and time (“news”). Al-Khalifa and Davis quantified the characteristics of del.icio.us folksonomy to determine prevalence of three classification categories: personal, factual, and subjective (Al-Khalifa and Davis 163). Personal tags are defined as self-referential and are often “used to organize a users own resources” (Al-Khalifa and Davis 164). Subjective tags assign value judgment. Factual tags “identify ‘facts’ about the described web resource” (Al-Khalifa and Davis 164). They concluded 34% were personal tags, 62% were factual tags, and 4% were subjective tags (Al-Khalifa and Davis 164), and therefore this distribution, indicated there was “meaningful semantics” in folksonomy tags (Al-Khalifa and Davis 166).

Others explored the semantic and structural aspects of folksonomy. Again, del.icio.us folksonomy tags are integral to this analysis. Certain semantic difficulties of folksonomy tags that confound users are polysemy (word or phrase with multiple meanings), synonymy, and “basic level variation”, which pertains to the inconsistency of describing an item in various instances as either general or specific (Golder and Huberman 200). Louise Spiteri drew upon

previous research to offer a comprehensive linguistic analysis of the structure of three folksonomy tagging environments: del.icio.us, Furl, and Technorati. Tags were evaluated using the National Information Standards Organization (NISO) guidelines for creating controlled vocabularies. Her results indicated that many tags to conform to these standards, but there are problems of ambiguous tags in the form of homographs and unqualified abbreviations and acronyms (Spiteri 21).

Technical Aspects of Implementing LibraryThing

LibraryThing assists with the initial set-up, but there are several things that your web OPAC administrator needs to accomplish before embedding the tags and tag browser widget into the relevant OPAC web pages. The administrator must create a tab delimited file containing ISBN (both 10 and 13 digit ISBNs are valid), title, and author of bibliographic records to upload to LibraryThing. A match is performed against LibraryThing.com's and the OPAC's ISBNs. If a record does not have a valid ISBN, then it will be ignored. It is not necessary to remove items that lack an ISBN. Once the data is sent, LibraryThing.com's technical support creates an institutional account to configure enhancements, widgets, etc. Once the administrator is notified that the match is complete, then he or she is permitted to configure the LibraryThing widget. The LibraryThing widget is a link to a JavaScript file that LibraryThing's technical support codes for each library requesting subscription service. Configuring the widgets does require some knowledge of Cascading Style Sheets (CSS) because libraries have access to the widget style sheet. In addition to configuring the look and feel of the tag browser, libraries also have control over how many tags display in their OPACs.

Installing LibraryThing tags and configuring the widget is a seamless process for most integrated library systems. In the rare instance of difficulty, there is a very active listserv, LibraryThing For Libraries, where fellow administrators and LibraryThing coders share their tips and answer questions.

Engaging Students in Social Tagging

After implementing a new service or feature, engaging your population is often the next challenge. Being familiar with the difficulties other libraries had in engaging their populations in generating folksonomy in the online catalog, we decided to target a particular population with an academic interest in the subject matter. The children's and young adult collections receive a great deal of use but we observed that the pre-service teaching students struggled in searching for books fitting the terminology they were using in their education classes. The Library of Congress Subject Headings are not very useful when a student is looking for many genres such as reality fiction or cumulative books. Additionally, the library had recently added a significant number of Spanish language children's books and there was concern about a lack of culturally relevant subject headings.

We approached three faculty from the College of Education with the idea of a service learning project for their students. The faculty members teach young adult literature, Spanish language arts and multicultural literature. We asked each to give their students an assignment to read pre-selected books from the two collections, and add educationally and culturally relevant social tags to titles in LibraryThing. They enthusiastically agreed realizing that discussion on social tagging could facilitate their students' understanding of the books and how they could be used in an educational setting.

The education librarian met with each class early in the semester. At each session, she demonstrated the library catalog, described the service-learning need, and provided instructions on adding social tags to LibraryThing. At the time of this writing, the multicultural literature class had not yet met. However, students in the young adult literature and Spanish language arts classes were enthusiastic about their participation although for different reasons. Students in the young adult literature class were intrigued with using a new technology. The students in the Spanish language class were appreciative of the fact that they could contribute culturally relevant terminology to the library catalog and were interested in bilingual tagging.

The number of books read and described by each student varied according to their class. Students in the young adult literature class were assigned two books each as they were reading longer texts. As the Spanish language class was reading shorter picture books, they read sixteen each. The professors required some books to be read by more than one person as a way of using the social tags to generate discussion in class.

Assigning Tags in LibraryThing

Even though LibraryThing tags are integrated into the online catalog, it is currently not possible for users to assign tags directly in the online catalog via the LibraryThing widget. The LibraryThing programmers are working on adding this feature. In the current iteration of the LibraryThing tag browser, the folksonomy will lead users to items similarly tagged in the library's collection. Users must create individual LibraryThing accounts and those tags will be included in the subsequent monthly tag import. It is also worth noting that only the top 20 tags are displayed in the OPAC. In the example below, the bibliographic record for Postcards from No-Man's Land is displayed. When a user selects a particular tag (in this instance, "amsterdam"), the widget displays as in the following example:

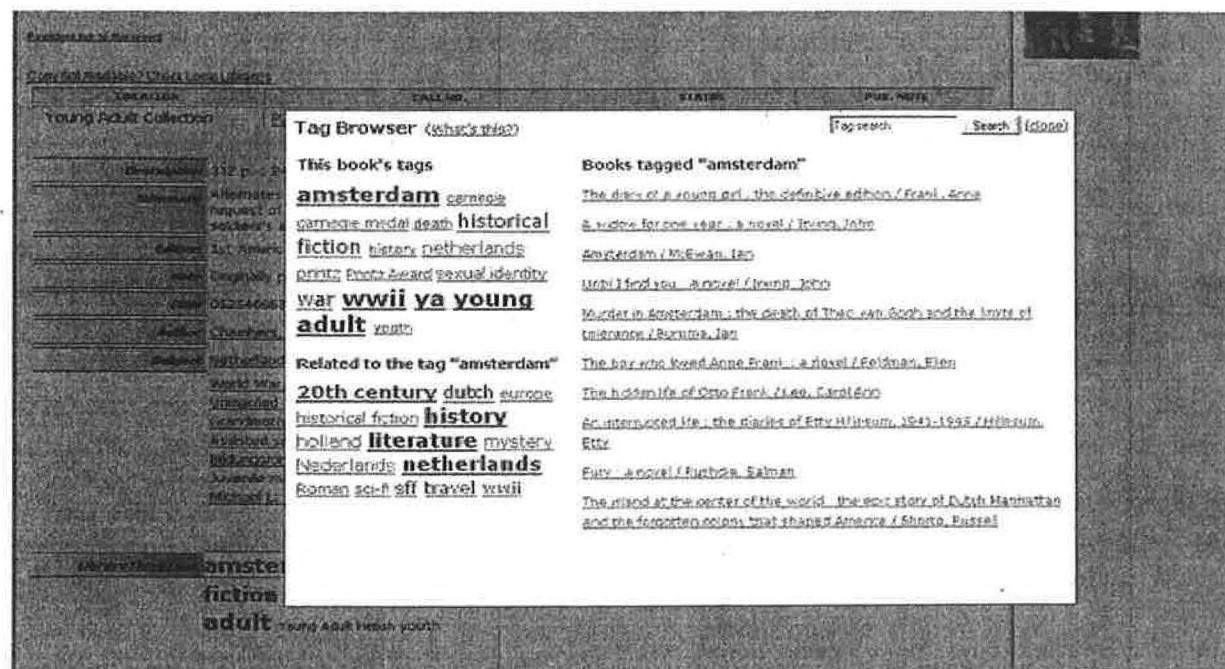


Fig.1. Display of LibraryThing Tag Browser.

Other tags related to the selected tag, “amsterdam”, are presented, offering serendipitous discovery of related books. For example, a user may decide to select the tag “netherlands” to see other items related to “amsterdam”. The Tag Browser also displays other books tagged “amsterdam” in the library’s collection. It is also possible to perform a tag search and the results yielded will show books in the library’s collection with that particular tag.

Analysis of User-Generated Folksonomy

In order to conduct a before-and-after comparison, we captured the social tags for each of the pre-selected books prior to meeting with the classes. One month after introducing LibraryThing to the students, we went back to the web site and recaptured the social tags a second time. In reviewing the new tags, we confirmed several of the findings from the research mentioned previously, as well as a few surprises (see Fig.2).

As Al-Khalifa and Davis found, factual tags, in this instance describing themes/genres and conveying subjects, were the most prevalent additions. They were also consistently very specific and supported Munk and Mork’s finding that those with professional expertise, such as teacher education, were less likely to use general tags. The tags revealed an awareness of using variant terminology for the same concept: loss of parent, parent death, and parent loss. Interestingly, tags for biographies were descriptive of the person and their life, and not necessarily describing the book. For example, a book about Muhammad Ali had tags “Draft dodger,” “Parkinson’s disease,” “Determination,” and “Heavyweight.” This made us wonder if the genre of the work affects the nature of the factual tags. Unfortunately, there was too little variety of genres for us to analyze this further.

We also saw some evaluative tags such as “excelente libro” and its English translation “great book”, but not as many as in LibraryThing as a whole. This last example also shows the Spanish language students’ active translation of preexisting tags for many books. Contrary to the aforementioned previous findings, there were very few self-referential tags. Occasionally, a book was tagged with “unowned” and “owned,” but this was very uncommon. Another surprise was the lack of curriculum-related tags except for an indication of the appropriate audience for a book. As these were teacher education students, we expected to see more classroom-use tags.

LibraryThing Folksonomy for <u>Diego</u>	
Pre-existing Tags	Tags Added by Students
3 rd right shelf	artista mexicano
Art	curandera inigena
Artists	Digeo Rivera
bilingual	historia de vida
biography	indigenous healer
heroes	knowledgeable
Level 0-blue	life history
Mexico	Mexican artist
my world and others	

Fig.2. Comparison of pre-existing tags to tags added by students.

Conclusion

This initial review of targeted social tagging was a fascinating introduction into what is possible with interactive catalogs. By combining social tagging with a service learning project, students were able understand the broader implications of their work and its affect on their local environment. For its part, the library received a richer catalog of new and meaningful access points for resource discovery.

This project indicates two future directions for further investigations. First, more refined analysis of the tags should be explored. Our discovery of how the nature of the factual tags is contingent on the genre of the book should be considered more fully. Additionally, the perspective of the student should also be considered. Future investigations should further explore the student experiences with social tagging in both creating them and later using the tags to locate relevant library resources.

Works Cited

- Al-Khalifa, H. S. and Hugh C. Davis. "Towards Better Understanding of Folksonomic Patterns." Proceedings of the 18th Conference on Hypertext and Hypermedia, September 10-12. ACM: New York, 2007. 163-166. ACM Portal. Criss Lib., Omaha, NE. June 9, 2008.
- Golder, Scott A., and Bernardo A. Huberman. "Usage Patterns of Collaborative Tagging Systems." Journal of Information Science 32.2 (2006): 198-208. Library Literature & Information Full-Text. WilsonWeb. Criss Lib., Omaha, NE. June 9, 2008.
- Munk, Timme Bisgaard, and Kristian Mørk. "Folksonomies, Tagging Communities, and Tagging Strategies--an Empirical Study." Knowledge Organization 34.3 (2007): 115-27. LISTA. EBSCOhost. Criss Lib., Omaha, NE. June 9, 2008.
- Spiteri, Louise F. "The Structure and Form of Folksonomy Tags: The Road to the Public Library Catalog." Information Technology and Libraries 26.3 (2007): 13-25. Library Literature & Information Full-Text. WilsonWeb. Criss Lib., Omaha, NE. June 9, 2008.
- Trant, J. "Exploring the Potential for Social Tagging and Folksonomy in Art Museums: Proof of Concept." New Review of Hypermedia & Multimedia 12.1 (2006): 83-105. LISTA. EBSCOhost. Criss Lib., Omaha, NE. June 9, 2008.