Tagging and Library Online Catalogs: An Annotated Bibliography

Debbie Farrar

Drexel University

Info 522

March 16, 2010
Tagging and Library Online Catalogs: An Annotated Bibliography

Introduction

Tagging, also known as social tagging and collaborative tagging, can be described as an interactive way for users to categorize and describe subject matter in order to give personal meaning to content and to possibly share that meaning with others. Tagging has been used with great success on the web for categorizing and describing subject matter at sites such as Library Thing, Delicious, and Amazon. Integrating tagging into libraries’ online catalogs would allow users to interact with other users through the use of tags and could possibly improve subject search success by linking user tags to Library of Congress subject headings is the topic of the following bibliography.

Description

The following bibliography includes articles suggesting that user supplied tags can enhance the use of library online catalogs and the possibility of linking tags to Library of Congress subject headings or the use of user tags alone could improve search results for users of library online catalogs. The articles in this bibliography cover research conducted on the topic of user tags within the last eight years from 2002 to 2009. All of the studies were conducted in the United States with the exception of one study conducted in New Zealand (Wetterstrom, 2008).


**Literature Review**

Subject searching in library online catalogs can often cause frustration for an untrained user. Users often use keywords and natural language to perform searches in library online catalogs as they do when searching for information on the web using search engines. In many cases the term used to search is not part of a controlled vocabulary used to index content so no results are received when searching library online catalogs. When a user does not receive results the user often draws a conclusion that no information on the topic exists in the library. On the other hand, if too many results are received, a user can easily be overwhelmed. Without the knowledge to refine a search a user will often resort to other means such as the web to locate needed information. Users today are accustomed to web search engines such as Google, Yahoo, Bing, and multitudes of other search engines that allow users to search for information using keywords and natural language (Antell, 2008; Halcoussis, 2002, Yu, 2004). In most, if not all cases, searches using web search engines do yield results. The fact that results are retrieved, even if the number of results is large, indicates to many users they have performed successful searches because results are usually ranked by relevance.

In order to make online library catalogs more user friendly, research has been conducted analyzing the tags users have placed on websites such as Library Thing and Delicious to see if there is a correlation between user supplied tags and Library of Congress subject headings (Rolla, 2009; Spiteri, 2007; Thomas, 2009; Wetterstrom, 2008; Yi, 2009). The results of these studies have indicated that in many cases user supplied tags often correlate with Library of Congress subject headings. Also,
researchers who conducted these studies found that where there was no direct correlation between the user tag and the subject heading, user tags can further refine subject headings.

I believe the studies regarding the correlation and linkage of user-supplied tags and Library of Congress subject headings are important as this research can be the groundwork for improving library online catalogs. The ability to link a tag that is user created to a subject heading could allow untrained catalog users to search the catalog using natural language searches. The linkage would be made in the background but the user would see results as if the search were done by a trained professional making the catalog and searching far more user-friendly.

Another aspect of tagging is that tags are often added to content not only for categorizing the content but also for personal reasons (Antell, 2008; Lawson, 2009; Golder, 2006). For example, a tag could be added to a book to describe it as “funny”. Other users can use a tag or keyword “funny” along with other tags or keywords to locate books that others found “funny” increasing satisfaction in results. While this type of tag would not specifically be considered a category, it can give the catalog value and enhance the user’s ability to search as well as make the catalog interactive.

There have also been several studies conducted where user tags have been implemented in online library catalogs and their uses has been analyzed (Mendes, 2009; Peterson, 2009). While in both cases, the use of user-supplied tags was low from the standpoint of using tags to search for information. The authors of the research, in both cases, also indicate the results may have been attributed to factors such as
Tagging and Library Online Catalogs: An Annotated Bibliography

placement on the screen used for searching and lack of communication to users about tagging. Another aspect I noted is that in both studies, the tags were not linked to Library of Congress subject headings and users who perform searches were not using tags to search the catalog but were observing the tags supplied to content found.

While research is still needed on the topic of tagging and ways to link tagging to currently indexed content in library catalogs, results of studies thus far are encouraging and a step toward making library online catalogs more user-friendly and interactive. Today users are expecting to have the ability to search for information in the library as they do on the web, user tagging offers libraries one way of meeting those expectations.

Bibliography

Entry 1

Abstract: Subject searching in the OPAC is the most problematic of all search types, causing far greater difficulty for patrons than keyword searching and known-item searching. This study combines two methodologies -- transaction log analysis and user observation interviews -- to examine the reasons for patrons' failure to use subject searching effectively. The transaction log analysis shows that patrons rarely utilize correct and complete subject terms and that they retrieve zero results in almost half of their searches. Furthermore, the user observation interviews reveal that users generally are unaware of the many tools and services that librarians have created to assist them with subject searching, and that asking a librarian for help simply does not occur to them. Even when searchers do locate and employ subject terms, the subject terms appear not to help them very much: Analysis of observed searches reveals almost no correlation between finding a subject term and judging a subject search to be successful. The authors discuss the potential for further research on "just-in-time" instruction, online instruction, and "tagging" as possible strategies to improve patrons' searching success.

Annotation: How users perform searches in online catalogs is studied and analyzed in this article. The study suggest that keyword searching is the most popular form of
searching catalogs and that searching subjects can be complimented with the implementation of user supported tags.

**Search Strategy:** I discovered descriptors for other articles I located in databases that I could possibly use to find articles pertaining to the topic of tagging and the possible use of tagging in catalogs. I also wanted to use multiple databases to search with keywords in Dialog.

**Database:** Dialog [Infosci]

**Method of Searching:** Keyword Searching

**Search String:**
- ? s online catalog
- S1 356 ONLINE CATALOG
- ? s tagging
- S2 23223 TAGGING
- ? s s1 and s2
- 356 S1
- 23223 S2
- S3 3 S1 AND S2

**Entry 2**

**Abstract:** Collaborative tagging describes the process by which many users add metadata in the form of keywords to shared content. Recently, collaborative tagging has grown in popularity on the web, on sites that allow users to tag bookmarks, photographs and other content. In this paper we analyze the structure of collaborative tagging systems as well as their dynamic aspects. Specifically, we discovered regularities in user activity, tag frequencies, kinds of tags used, bursts of popularity in bookmarking and a remarkable stability in the relative proportions of tags within a given URL. We also present a dynamic model of collaborative tagging that predicts these stable patterns and relates them to imitation and shared knowledge.

**Annotation:** The study performed in this article indicates there is consistency in how users tag content. It also suggests that while most tagging is done for personal reasons, those tags can benefit others.

**Search Strategy:** This article was referenced in Karen Lawson’s article, “Mining Social Tagging Data for Enhanced Subject Access for Readers and Researchers”. I did a title search in the LISTA database to locate the full text of the article.
Database: LISTA

Method of Searching: Footnote chasing

Search String: TI Usage patterns of collaborative tagging systems

Entry 3


Abstract: Data from an observation study of a Web catalog in a small private arts college library are used to analyze the determinants of user success and satisfaction. Multiple regression models are estimated to identify the most important causative factors determining catalog user success in finding information, user attitudes to catalog organization, and user ability to navigate the catalog. It is found that subject-search users are more likely to assign a low score to catalog organization and to encounter difficulty navigating the catalog than users of known item and other search methods. These findings accord with the extensive literature on the problems associated with subject searching. Also, it is found that the more time spent searching and the larger the number of search results, the more likely it is that the user would report difficulty navigating the catalog. A significant result is that although the user’s perception of success or failure of the search is the most important factor determining both the user’s evaluation of the catalog organization and the navigability of the catalog, the success or failure of the search itself is not explained by any other variables included in the model. This exogeneity of search success has important implications for library instruction because it suggests that a user’s perception of success is dependent on the expectations the user brings to the search rather than specific features of the catalog design.

Annotation: In this study, the authors, using multiple regression models, have found that subject searching in online catalogs are the least successful from the user’s point and view.

Search Strategy: Reference to this article was made in the article by K. Antell. In order to find the article I searched for the journal title “Information Technology and Libraries” using the journal search feature at Hagerty and located the article using the citation information.

Method of Searing: Footnote Chasing

Search String: Full Text: Wilson Library Lit and Information Science
Year: 2002 Volume: 21 Issue: 4 Start Page: 148
Entry 4

**Abstract:** Social tagging enables librarians to partner with users to provide enhanced subject access. This paper quantifies and compares LC subject headings from each of 31 different subject divisions with user tags from Amazon.com and LibraryThing assigned to the same titles. The intersection and integration of these schemas is described and evaluated.

**Annotation:** The article suggests that tagging can enhance subject access to catalogs. The author further suggests that user supplied tags can make online catalogs more inclusive and helpful to users.

**Search Strategy:** I used the Research Guides provided on the Hagerty Library website to find articles related to topics in Library and Information Science. I was doing my initial search on the topic and used keyword searching in the database.

**Database:** LISTA

**Method of Searching:** Keyword searching

**Search String:** LibraryThing AND Catalogs

Entry 5

**Abstract:** The purpose of this paper is to present the implementation of LibraryThing for Libraries (LTFL) in an academic library and analysis of usage of LTFL data and their potential for resource discovery in the catalog.

**Annotation:** The article analyzed the implementation of the LibraryThing for Libraries (LTFL) into the catalog at Oviatt Library at California State University at Northridge and studied the usage statistics provided by LTFL. The authors concluded the usage of LTFL information by users was low, in the 1 to 2 percent range. The authors further concluded the low usage may have been the result of where the information was located, the addition of LTFL was not communicated to the users, and the study was performed at a generally low-usage period in the library.

**Search Strategy:** Initial search in dialog to discover articles that pertain to the chosen topic. I used keyword searching to locate articles.

**Database:** Dialog [File 438: Library Literature & Information Science]

**Method of Searching:** Keyword Searching
**Entry 6**

Peterson, E. (2009). Patron preferences for folksonomy tags: Research findings when both hierarchial subject headings and folksonomy tags are used. *Evidence Based Library and Information Practice, 4*(1), 53-56.

**Abstract:** With the emergence of folksonomy as an option for subject tagging, discussions have ensued about the costs and benefits of continuing to construct and apply traditional subject headings, given that patrons now can generate their own tags. To date, there are very few databases that allow both systems to coexist. Within the full text ETD (Electronic Theses and Dissertations) database at Montana State University, both traditional, hierarchical subject headings and patron applied tags are allowed. Patrons are encouraged to tag and the database even features a browse tag capability and a featured ETD. After 24 months of coexistence, data was gathered and analyzed to determine patron use and preferences when given the option of adding their own tags. Adapted from the source document.

**Annotation:** This study looked at the possibility of replacing Library of Congress subject headings with user supplied tags in order to reduce the cost of cataloging dissertations and thesis in a special database at Montana State University. The study found while patrons did tag documents, the tags supplied could not replace the Library of Congress subject headings.

**Search Strategy:** I wanted to try a controlled vocabulary search in LISA as part of my initial search on my topic.

**Database:** LISA

**Method of Searching:** Controlled Vocabulary search

**Search String:** (DE="tagging") and(DE=("cataloguing" or "bibliographic records" or "catalogues" or "cataloguing departments" or "indexing" or "tagging"))

**Entry 7**

Abstract: Some members of the library community, including the Library of Congress Working Group on the Future of Bibliographic Control, have suggested that libraries should open up their catalogs to allow users to add descriptive tags to the bibliographic data in catalog records. The website LibraryThing currently permits its members to add such user tags to its records for books and therefore provides a useful resource to contrast with library bibliographic records. A comparison between the LibraryThing tags for a group of books and the library-supplied subject headings for the same books shows that users and catalogers approach these descriptors very differently. Because of these differences, user tags can enhance, subject access to library materials, but they cannot entirely replace controlled vocabularies such as the Library of Congress subject headings.

Annotation: The article is a study comparing the use of LibraryThing tags and Library of Congress subject headings and concluded that while user supplied-tags can enhance catalog records, they cannot replace controlled vocabularies. The article illustrated user-supplied tags produce far more subject-type headings in records compared to subject headings in Library of Congress subject heading in records. The article further suggests user-supplied tags may be more useful in public library settings than in academic library settings.

Search Strategy: Initial search in dialog to discover articles that pertain to the chosen topic. I used keyword searching to locate articles.

Database: Dialog [File 438: Library Literature & Information Science]

Method of Searching: Keyword Searching

Search String: ? b438
? s tagging
   S1  95 TAGGING
? s library and s1
   109689 LIBRARY
   95 S1
   S2  25 LIBRARY AND S1

Entry 8

Abstract: This article examines the linguistic structure of folksonomy tags collected over a thirty-day period from the daily tag logs of Del.icio.us, Furl, and Technorati. The tags were evaluated against the National Information Standards Organization (NISO) guidelines for the construction of controlled vocabularies. The results indicate that the tags correspond closely to the NISO guidelines pertaining to types of concepts expressed, the predominance of single terms and nouns, and the use of recognized spelling. Problem areas pertain to the inconsistent use of count nouns and the incidence
of ambiguous tags in the form of homographs, abbreviations, and acronyms. With the addition of guidelines to the construction of unambiguous tags and links to useful external reference sources, folksonomies could serve as a powerful, flexible tool for increasing the user-friendliness and interactivity of public library catalogs, and also may be useful for encouraging other activities, such as informal online communities of readers and user-driven readers’ advisory services.

**Annotation:** The article is the results of a study performed to determine how closely user tags relate to standard controlled vocabularies and concluded that user tags do closely correspond to controlled vocabulary terms. The article also pointed out advantages and weaknesses of user tags but suggests value can be found in the use of user tags.

**Search Strategy:** Initial search in dialog to discover articles that pertain to the chosen topic. I used keyword searching to locate articles.

**Database:** Dialog [File 438: Library Literature & Information Science]

**Method of Searching:** Keyword Searching

**Search String:** ? b438
? s tagging
   S1 95 TAGGING
? s library and s1
   109689 LIBRARY
      95 S1
   S2 25 LIBRARY AND S1

**Entry 9**

**Abstract:** The purpose of this article is to provide a quantitative analysis of the extent to which folksonomies replicate the Library of Congress Subject Headings (LCSH) to see if folksonomies would successfully complement cataloger-supplied subject headings in library catalogs. Design/methodology/approach--The paper compares social tags and LC subject headings for ten books from various library-related applications including next generation OPACs and LibraryThing by ranking tags and subject headings using scales modified from research by Golder and Huberman, Voorbij, and Kipp. Findings--Social tagging does indeed augment LCSH by providing additional access to resources. Research limitations/implications - Several of our applications lacked tags for the books we chose in our study. Tags are primarily taken from LibraryThing. Practical implications--A hybrid catalog combining both LCSH and a folksonomy would result in richer metadata and be stronger than the sum of its parts, giving patrons the best of both worlds in terms of access to materials. Originality/value--This paper supplies quantitative support for the use of folksonomies in a library’s
catalog. The data also supports many of the previous theories proposed in literature about folksonomies and social tagging.

**Annotation:** This article concerns itself with an extensive study comparing user supplied tags to Library of Congress subject headings and controlled vocabulary. The results of the study confirmed that many user supplied tags closely match those of controlled vocabulary. The study also suggests that tags that do not fall in line with controlled vocabulary can be used to further enhance online catalogs.

**Search Strategy:** I used the Research Guides provided on the Hagerty Library website to find articles related to topics in Library and Information Science. I was doing addition research on the topic and used keyword searching in the database.

**Database:** LISTA

**Method of Searching:** Keyword search

**Search String:** folksonomy AND catalog

**Entry 10**

**Abstract:** This study investigated the complementarity of user-assigned tags and Library of Congress Subject Headings (LCSH) assigned by cataloguers in a New Zealand library context. In order to identify the added value of tags in a library catalogue, the study experiment asked 20 adult New Zealanders to tag between 9 and 15 books from the general collection of the National Library of New Zealand. A total of 897 tags were allocated to 217 books. Seventy five percent of the tags did not match any LCSH, which indicates a high level of tag complementarity. The experiment also showed that the majority of the non-matching tags were either of a more popular language (21.63%) or indicated a different point of view through a related term (19.29%). Different levels of specificity were also common: 14.16% of the tags were broader terms, and 19.62% were narrower in scope than the LCSH. The results did not show significant added value in regards to New Zealand English vocabulary, or concerning currency of the terms.

The study also identified indications of collaborative value. Tags that had been allocated to books about similar topics were grouped together to form tag-clouds. The average number of allocated tags within each tag-cloud was 24.4 tags, and an average of 2.9 of these tags, or 11.9%, were shared by two or more people. However, the vast majority of the shared tags were only shared by two people, while a higher degree of collaboration was relatively rare.
Annotation: The article suggests that tagging can enhance library online catalogs especially in removing barriers that Library of Congress subject headings may impose because of language differences. The study was conducted in New Zealand and demonstrated how terms used in Library of Congress subject headings may differ from those used in everyday vocabulary.

Search Strategy: I wanted to try a controlled vocabulary search using Dialog as part of my initial search on my topic.

Database: Dialog [File 438: Library Literature & Information Science]

Method of Searching: Controlled Vocabulary search

Search String:  
? s tagging (internet)  
S4 54 TAGGING (INTERNET)  
? s lcsh  
S5 118 LCSH  
? s s4 AND s5  
54 S4  
118 S5  
S6 1 S4 AND S5

Entry 11  

Abstract: The purpose of this paper is to investigate the linking of a folksonomy (user vocabulary) and LCSH (controlled vocabulary) on the basis of word matching, for the potential use of LCSH in bringing order to folksonomies. Design/methodology/approach -- A selected sample of a folksonomy from a popular collaborative tagging system, Delicious, was word-matched with LCSH. LCSH was transformed into a tree structure called an LCSH tree for the matching. A close examination was conducted on the characteristics of folksonomies, the overlap of folksonomies with LCSH, and the distribution of folksonomies over the LCSH tree. Findings -- The experimental results showed that the total proportion of tags being matched with LC subject headings constituted approximately two-thirds of all tags involved, with an additional 10 percent of the remaining tags having potential matches. A number of barriers for the linking as well as two areas in need of improving the matching are identified and described. Three important tag distribution patterns over the LCSH tree were identified and supported: skewedness, multifacet, and Zipfian-pattern. Research limitations/implications -- The results of the study can be adopted for the development of innovative methods of mapping between folksonomy and LCSH, which directly contributes to effective access and retrieval of tagged web resources and to the integration of multiple information repositories based on the two vocabularies. Practical implications -- The linking of controlled vocabularies can be applicable to enhance information retrieval capability within collaborative tagging systems as well as across various tagging system
Tagging and Library Online Catalogs: An Annotated Bibliography

information depositories and bibliographic databases. Originality/value -- This is among frontier works that examines the potential of linking a folksonomy, extracted from a collaborative tagging system, to an authority-maintained subject heading system. It provides exploratory data to support further advanced mapping methods for linking the two vocabularies. Adapted from the source document.

Annotation: This article studies the possibility of linking user created tags to Library of Congress subject headings, matching uncontrolled vocabulary to controlled vocabulary in order to provide structure. The authors further suggest that linking could further enhance information retrieval in online catalogs.

Search Strategy: I wanted to try a controlled vocabulary search in LISA as part of my initial search on my topic.

Database: LISA

Method of Searching: Controlled Vocabulary search

Search String: (DE="tagging") and(DE=("cataloguing" or "bibliographic records" or "catalogues" or "cataloguing departments" or "indexing" or "tagging"))

Entry 12

Abstract: This paper analyzes the results of transaction logs at California State University, Los Angeles (CSULA) and studies the effects of implementing a Web-based OPAC along with interface changes. The authors find that user success in subject searching remains problematic. A major increase in the frequency of searches that would have been more successful in resources other than the library catalog is noted over the time period 2000-2002. The authors attribute this increase to the prevalence of Web search engines and suggest that metasearching, relevance-ranked results, and relevance feedback ("more like this") are now expected in user searching and should be integrated into online catalogs as search options.

Annotation: This article examines the way users search for information in online catalogs and suggests changes in online catalog interfaces may enhance the success rate of user searches. Suggestions include the use of keyword searches using natural language linked to controlled vocabulary using system maps, relevance-ranked results, and feedback.

Search Strategy: This article was referenced in K. Antell and J. Huang’s article “Subject searching success: Transaction logs, patron perceptions, and implications for library instruction”. I did a title search in the LISTA database to locate the full text of the article.
Database: LISTA

Method of Searching: Footnote chasing

Search String: TI Transaction logs, patron perceptions, and implications