

User Tags versus Subject Headings

Can User-Supplied Data Improve Subject Access to Library Collections?

By Peter J. Rolla

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.

The advent of interactive websites, part of what is known as the Web 2.0 or second-generation Web development and design, has called into question the ways in which libraries provide access to their collections. Today's library users, who are increasingly comfortable with searching on the Internet, have certain expectations about how to search for information and how it will be displayed. These expectations, however, do not match how information is contained, discovered, and presented in traditional library catalogs. A recent study, for example, found that students using the University of Oklahoma's online public access catalog (OPAC) performed keyword searches fourteen times more often than subject searches.¹ In addition to a reliance on keyword searching, today's users increasingly use interactive websites that allow them to both upload their own data or content and to connect with other users of the site—the Web 2.0 phenomenon. Facebook, MySpace, and YouTube are several currently prominent examples of websites that thrive on user-supplied content, but even a now venerable site like Amazon has always allowed its customers to post reviews and comments. This paper will look closely at LibraryThing (www.librarything.com), a website that could be considered a Web 2.0 version of a union catalog.

Many of today's most popular websites allow users to "tag" specific content; that is, users can supply their own keywords to describe websites, images, or other content. User-supplied tags of this type potentially offer a way for libraries to improve subject access to the materials in their collections. The Library of Congress (LC) Working Group on the Future of Bibliographic Control, convened to study the present state of cataloging in libraries and to make recommendations

Peter J. Rolla (peter.rolla@colorado.edu) is Cataloging Librarian at University of Colorado at Boulder.

An earlier version of the paper was presented at the Third Colorado Academic Library Summit, May 23, 2008 in Lakewood, Colorado, under the title "LibraryThing vs. the Library Catalog."

Submitted October 8, 2008; returned to author December 12, 2008 with request for revision and resubmission; revised and resubmitted February 27, 2009, and reviewed and accepted for publication March 28, 2009.

for the future, recommended that libraries allow users to add tags and other user-supplied data to their catalogs.² The LC Working Group noted that allowing user-supplied data in online catalogs will make the catalogs more relevant to users accustomed to the Internet and also will improve access to the materials in library collections. Most libraries currently provide subject access to their materials through Library of Congress subject headings (LCSH) supplied by catalogers. Professional catalogers typically perform this work, since LCSH are governed by a complicated set of rules that requires training and specialized knowledge to follow. Providing subject access to collections, therefore, is an expensive part of cataloging work, since it is time-consuming and usually performed by professional staff. In addition, in an environment where users are accustomed to keyword searches on the Internet, many librarians question the value of the complicated pre-coordinated subject strings that make up an LCSH. The Working Group observed that “the creation of pre-coordinated subject strings, combining the topical, geographical, chronological, and genre aspects of a work into a single subject heading, can be a time-consuming and complex process. . . . While pre-coordination can offer users an implicit indication of the relationship between subject terms, the carefully crafted subject strings created by catalogers are often misunderstood or incomprehensible to users and reference librarians.”³ In recommendation 4.1.2, the Working Group explicitly advises integrating user-contributed data into library catalogs.

Many technological and policy issues are involved in opening online catalogs to user-supplied data, but the present study will address even more fundamental questions: How do user tags differ from the subject headings assigned by catalogers? Will user tags provide better subject access than the cataloger-supplied subject headings? Can user tags provide insight into how readers think about the subjects of books and therefore suggest ways in which library-supplied subject access can be improved? An exploratory and initial comparison of tags and subject headings will lay the groundwork for further research.

This paper will use the term *tags* to refer to the descriptors, which may be single words or phrases, assigned to a website or other resource, typically by the users of the site. *Folksonomy* refers to the collective grouping of tags assigned by an aggregate of users of a particular website. Finally, the term *tag cloud* refers to the display of tags using visual cues, like size, color, and proximity to indicate the importance of terms or their relations to each other.

Literature Review

During the past several years, articles that discuss user tags and folksonomies have begun appearing in library and

information science journals. The appearance and popularity of Web 2.0 sites like Delicious (<http://delicious.com>) and Flickr (www.flickr.com), which allow users to add tags to the sites' content, have inspired researchers to take a look at these tags and to explore how libraries might incorporate user tags into their Web-based services, including, but not limited to, the online public catalog.

Most of the authors who have studied tags and folksonomies generally feel that user tagging would enhance libraries' websites and catalogs. Spiteri, for example, believes that allowing user tags on a library site can supplement controlled vocabularies.⁴ User tags would also permit patrons to personalize the library's website, thereby bolstering a spirit of belonging and also fostering online communities organized around the library. Fichter agrees that tags, since they are popular and fun to work with, can help users feel more connected to the library's website.⁵ In her opinion, tags have a low barrier to participation because users do not need to learn complicated thesauri or controlled vocabularies. Once users breach this low barrier and see the ease and personal benefits of tagging, the social aspect then encourages more tagging. Fichter calls user tags “nimble and flexible,” a sentiment echoed by Spiteri.⁶ Spiteri notes that folksonomies can more easily accommodate new terms and concepts than heavily controlled vocabularies like LCSH.⁷ Fichter and Spiteri also point out that controlled vocabularies often do not use natural language, and therefore tags more closely represent how readers think and speak about a subject.

Although many of the authors who have studied tags and folksonomies generally display a positive attitude toward them, some recognize the inherent weaknesses of and problems with user tags. Golder and Huberman, for example, point out that tags, unlike controlled vocabularies, do not deal with problems created by polysemy (one word having multiple meanings), synonymy (more than one word with the same or similar meanings), and basic-level variation.⁸ The latter term refers to the continuum of meaning, from general to specific, that potentially exists for any given concept. To explain the concept of basic-level variation, Golder and Huberman give the example of a cheetah, which could be assigned various subject terms from most specific to most general: a cheetah, a cat, or an animal. In a completely uncontrolled, user-driven folksonomy, no rules exist to govern the level of specificity when assigning terms. Likewise, user tags do nothing to solve the problems of polysemy and synonymy, whereas one of the main purposes of controlled vocabularies is to disambiguate polysemous words and choose preferred terms from groups of synonyms. Golder and Huberman also bring up a key feature of user tags, one that is readily apparent in the tags on LibraryThing. Although tagging does allow the use of terms that are helpful to the community as a whole (i.e., terms that would allow other users to discover a certain website or book), it also

permits the inclusion of terms that are personal in nature and only helpful to the person adding the term.

Other researchers who have studied user tags have discovered features of folksonomies that, while not inherently positive or negative, should be considered if libraries allow users to add tags to libraries' websites or catalogs. Munk and Mørk performed a statistical study of more than seventy thousand keyword tags on Delicious and found that the keywords assigned to a specific resource (websites, in the case of Delicious) do form a distinct pattern.⁹ According to their study, only a very few keywords dominate the group of tags assigned to a resource. In their words, "These [few] keywords are primarily the so-called cognitive basic categories and essentially consist of a number of very broad and general content categories that are common to all people."¹⁰ Users tend to pick terms representing the broader end of the basic-level variation continuum. Golder and Huberman also found this to be true.¹¹ Munk and Mørk do find that the dominance of a few broad keywords potentially minimizes the usefulness of tags, since broad and general keywords do not enhance surprise and discovery. Another noteworthy feature of user tags that several researchers noted involves not the words that make up the tags themselves but rather who inputs tags. Munk and Mørk, for example, when looking at Delicious, found that many of the popular keywords were related to computers and information technology (IT).¹² These IT-related tags also had a level of specificity that did not quite conform to the pattern mentioned above, in which general, broad concepts dominated the popular user tags. The prevalence of computer and IT-related tags indicates that a large portion of Delicious's users work in or have a strong interest in computer-related fields. The fact that a disproportionate number of keywords within a particular folksonomy involve a specific discipline suggests a potential problem for libraries who want to allow user tags. If a library serves a more general population overall, it does need to be aware that user tags may come predominantly from specific populations or communities.

The research on user tags to date has looked primarily at tags that describe Web-based or digital resources. Delicious, for example, which was one of the first popular websites to allow tagging, has been the subject of several of the studies discussed here. Some researchers have also looked at tags on Flickr, a photo-sharing site that allows its users to assign tags to the images they upload.¹³ The present study, however, will consider the tags on LibraryThing, a website that allows users to assign tags to books. This study will compare the user tags for a set of books on LibraryThing with the LCSH assigned to the same set of books. Wetterstrom conducted similar research, comparing user tags to LCSH.¹⁴ Wetterstrom asked a group of twenty volunteers to come up with tags for a collection of books and compared these tags to LCSH assigned to the same materials. The present

study, which uses LibraryThing tags as a basis of comparison, complements Wetterstrom's work, since the tags on LibraryThing were created by a significantly larger group of people than Wetterstrom's. Wetterstrom's specific conclusions about user tags and LCSH will be discussed later, since they differ from the present study's findings.

LibraryThing

Although library catalogs increasingly contain metadata about digital objects, websites, and other nonprint materials, libraries originally created their catalogs to describe and provide access to books and other printed matter. These materials still make up a large portion of what is represented in library catalogs. LibraryThing, then, more so than other sites that allow tagging, provides a useful comparison to the traditional library catalog and also can provide an example of what bibliographic records might look like if library catalogs are opened to user input. Like a library catalog, LibraryThing contains a comprehensive list of books, but, like other Web 2.0 sites, it allows users to interact with the content and supply their own data. Users of LibraryThing can create their own virtual libraries, rate books, and interact with other readers on the site. Also, and most importantly for this study, users can supply their own tags for their books. The integrated library systems (ILS) that currently run most library catalogs do not yet allow users to contribute data to bibliographic records. Some libraries already have begun to experiment with ways to incorporate tags and other user-supplied content into catalogs. The new generation of OPACs like Endeca, AquaBrowser, and Encore also offer tag clouds of various types. User tags, however, are still rare in a library environment, making this an excellent time to study whether they will help provide better subject access to library collections.

LibraryThing and Tagging

A useful first step is looking at LibraryThing's explanation of user tags and its instructions on how the site's members can apply them. In response to the question "What are tags?" the following answer appears on LibraryThing: "The short answer: Tags are a simple way to categorize books according to how you think of them, not how some official librarian does."¹⁵ Two aspects of this definition are worth noting: the openness of tags—users can categorize their books any way that is useful to them—and the fact that LibraryThing is placing itself in opposition to "official librarians." The site's "long answer" continues these two themes:

Once you have a hundred books or so, you need some way to organize them. Library subject

classifications, including that of the Library of Congress, are one solution. For most personal libraries, however, they aren't much use. "Tags," informal, personal markers used on blogs and sites like Flickr and Del.icio.us, provide a better model.

Here are two examples from my (Tim's) experience:

1. The LC catalogs Ben's *Aegean Turkey*, a guide to the archaeological sites of Turkey's western coast, under the single subject, "Ionia." For me, however, the book is about turkey [sic] and archaeology, tags I've applied to dozens of books, including Ben's other archaeological guides.
2. The LC thinks Bernadette Brooten's *Love between women: early Christian responses to female homosexuality* is about six different things, including the mouthful "Bible. N.T. Romans I, 18–32—Criticism, interpretation, etc.—History—Early church, ca. 30–600." I get by with the tags early church, and homosexuality. To these I added the tag divination. Although the book doesn't say much about divination, its comments on the topic were actually the reason I picked it up.

Tags can also mark "favorites" or "books to read." I've used the tag ben's to mark books I should return to my friend Ben. (That I included them in my catalog is, however, a bad sign for that!)¹⁶

These instructions clearly exhort the site's members to add tags for whatever reason they find useful, even for very personal instances like "books borrowed from Ben" and "books at the summer house." In addition, both sets of instructions place user tags squarely in competition with LCSH and make the claim that tags are better. The present study aims to explore that claim.

Research Method

Since this study is meant as an initial foray into this arena and not the definitive answer on tagging and subject access, it initially examined a small number of books. A sampling was required that was large enough to provide enough data to analyze but also small enough that the language used in individual tags and subject headings could be studied in detail. Having a sampling of books on a wide variety of subjects and dealing with a variety of geographic regions also was desirable. To accomplish these goals, three searches were performed in LibraryThing using its Tagmash feature (essentially a keyword search of the user tag field): one for

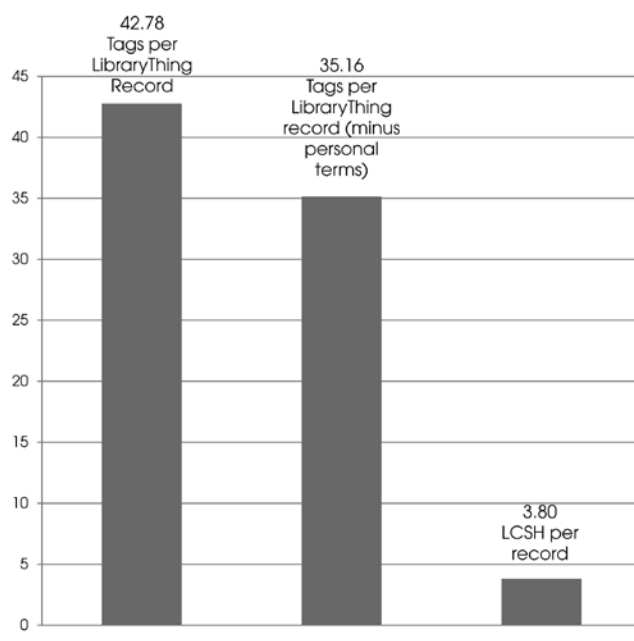


Figure 1. Average Number of Tags or Subject Headings per Record

the "nonfiction," one for "Africa" and "history," and finally one for "Mexico" and "immigration." The first fifteen titles returned on each of these searches, representing the titles in which the search terms were mostly frequently used, were chosen for the study. All of the titles chosen were in English. Choosing titles from these three searches accomplished the goal of having a sample that represented a variety of subject areas, and the author felt that the overall sample of forty-five books was an appropriate number for a small-scale and initial study. Once the list of books from LibraryThing was complete, these same titles were searched in OCLC's WorldCat, and then the user tags and LCSH for each title were compared. A list of the titles studied appears in the appendix.

Findings and Discussion

Numerical Comparison

Perhaps the most dramatic difference between the application of user tags and Library-assigned LCSH is that the website's users assign many more tags to books than library catalogers assign subject headings (see figure 1). Each of the forty-five books under consideration had more user tags in LibraryThing than subject headings in the catalog record by a large margin. The LibraryThing records for these titles had an average of 42.78 tags, while the library records had an average of 3.80 subject headings per record. In LibraryThing records, an average of approximately 7 tags

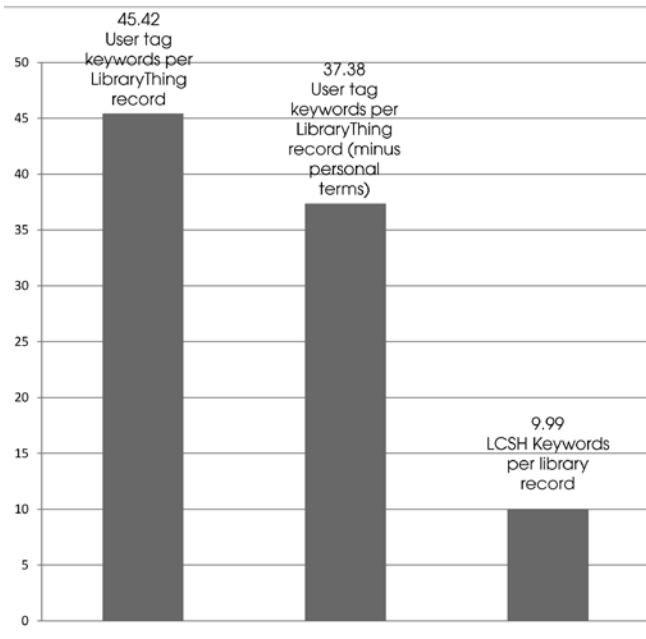


Figure 2. Average Keywords per Record

on each record consisted of personal terms; personal terms are explored in depth in the next section. Even disregarding these personal terms, there was an average of 35.16 user tags per record, still much higher than the average of 3.80 LCSH per record.

LibraryThing allows tags of more than one word, and LCSH are made up of strings of pre-coordinated terms, so both tags and subject headings can be broken down into individual keywords. Thus a tag cloud in LibraryThing might include the tags “British empire” and “British history,” which would count as two of the 42.78 average tags but would also count as three keywords, since the word “British” is repeated. Similarly, an LCSH subject string like “Mexico—Emigration and immigration—Social aspects” has five keywords in it, ignoring the conjunction. If keywords (not full tags or subject headings as seen in figure 1) are counted, then LibraryThing still averages more per record: an average of 45.42 per record (37.38, if the personal terms are not included) versus an average of 9.99 keywords represented by LCSH.

Personal Tags

Another aspect of LibraryThing tags is immediately noticeable. The user tags in these records contain many personal or individual terms, just as Golder and Huberman remarked in their study.¹⁷ Tonkin and colleagues found that personal terms, which they call “time, task, or project labels,” form 16 percent of unique tags on Delicious.¹⁸ LibraryThing’s instructions for tagging books encourage the use of personal terms. In addition to the examples of “Ben’s books” and

“summer house” found in the website’s instructions, terms like “book group,” “book club,” “read,” “unread,” “to read,” “own,” “read in 2007,” and “not in library” frequently appear in the records studied. Several tags of this type appear in each of the forty-five records studied here. These terms may have strong personal value to the users who input them but, from a library’s perspective, they are not useful descriptor terms and do not provide any subject access to the books. LibraryThing’s tag cloud display, like most such displays, does give weight to more popular terms by increasing the size of words in the display according to how many users have assigned those tags to a work. These personal or individual terms are usually very small, indicating that only a few people have added that tag to a record, so that visually, at least, these personal terms do not predominate in the tag cloud display. If libraries permit user tags in their catalogs, they will have to decide whether to allow and how to handle these individual and perhaps unhelpful tags. Conversely, certain personal tags (i.e., those that do not directly relate to the contents of the book but to the user’s experience with the book) could have practical value in a library catalog environment. For example, if students or professors come across books that are useful for a particular course, they could tag the book’s record with the course number to help other students find the same book.

Comparing User Tags and Subject Headings

In every LibraryThing record, the user tags contained at least one concept not covered by the subject headings in the catalog record. In many cases, these concepts represented ideas that a cataloger would not have brought out, deeming them irrelevant to the overall content of the book or somehow not consistent with the typical practice of subject analysis for books. Comparing LibraryThing’s instructions on assigning tags with the LC’s instructions to catalogers on how to assign subject headings illuminates some of the differences between the use of tags and subject headings. In section H180 of the *Subject Cataloging Manuals*, the LC instructs catalogers to “assign to the work being cataloged one or more subject headings that best summarize the overall contents of the work and provide access to its most important topics.”¹⁹ Compare this to the longer instructions given by LibraryThing, in which a reader assigned the tag “divination” to a work not because the book was primarily about that topic but because its comments on divination interested him. The fact that the users of LibraryThing assign tags to books representing concepts not brought out by LCSH does indicate that catalogers, by following the LC guidelines, may omit concepts that are important to users.

For each of the forty-five titles in this sample, the LibraryThing tags contained subject terms or concepts that the subject headings did not express. That figure does not

A savage war of peace : Algeria 1954-1962
by **Alistair Horne**

☆☆☆☆☆
(not yet rated)

Type: Book; English

Publisher: London : Papermac, 1987.

Edition: Rev. ed | [2 Editions](#)

ISBN: 0333413989 9780333413982

OCLC: 59096665

Related Subjects: [Algeria -- History -- Revolution, 1954-1962.](#) | [Algeria -- History -- Revolution, 1954-1962.](#)

Figure 3. WorldCat Record for *A Savage War of Peace: Algeria 1954-1962*

A Savage War of Peace: Algeria 1954-1962 (New York Review Books Classics)

by Alistair Horne

Add to your library

Members	Reviews	Popularity	Average rating	Conversations
268	6	16,307	☆☆☆☆ (4.24)	8

Members all members

Recently added by: Miro, cowleyleibrary, jensenmk82, yxf7h, braungrayson, kc324s, mack1968, wcflyfish, zenosbooks, debweiss

Member tags numbers | all tags

1950s 1960s 20th century @ **africa algeria** Algeria-History-Revolution 1954-1962 Algerian history Algerian Revolution Algerian War Algerian war Amazon Wishlist arab books to read COLD WAR cold war era **colonialism** counterinsurgency decolonization empire europe **france** french french history Guerrilla War **guerrilla warfare history** imperialism insurgency Islam liberation Maghreb middle east military **Military history** nationalism **Non-fiction North Africa nyrb** Paperback postwar pre-1983 revolution **terrorism** torture unread **War** western civilization

Figure 4. LibraryThing Record for *A Savage War of Peace: Algeria 1954-1962*

include the personal or individual terms, but words and phrases describing the subject of the book. Conversely, the librarian-assigned subject headings in twenty-five records (55.6 percent) brought out concepts and topics that the user tags did not. Finally, the subject headings and user tags assigned to thirty-five records (75.6 percent) brought out the same subject or concept, although often expressed in different terms. Thus, approximately three-quarters of the time, catalogers and readers agree on at least part of what each book is about, even if the tags and subject headings express the content of the book differently. The specific ways in which the user tags and subject headings for these titles differ is instructive, and these differences do suggest that adding user tags to library catalogs could help improve subject access to collections.

First, user tags almost always include very general and broad subject terms. In each of the forty-five records under consideration, LibraryThing's users added general or

broader terms for the concepts discussed. Books about Africa show this clearly. Several of the books under consideration discuss the genocide in Rwanda, and at least one deals with the civil war in the Congo. The catalogers, following the instructions for assigning LCSH, assigned headings relevant to the specific events and countries. For example, the subject heading string "Rwanda—History—Civil War, 1990–1993" appears in a bibliographic record. The LibraryThing records for books on Rwanda and the Congo, however, all contained tags such as "Africa," "African history," and other similar broader terms. Catalogers did not—and as a rule do not—include broader geographic terms. The books on Mexican immigrants show the same pattern. Catalogers used headings that included the term "Mexican American," but LibraryThing's users added terms like "Hispanic" and "Latino," which represent the broader population of which Mexican Americans are a part. Other examples from LibraryThing include the tags "science," "modern history," or "world history." Catalogers would only use a term like "science" for a book that is an introductory textbook or that discusses the entire discipline of science. Similarly, a book with a cataloger-supplied subject heading "World history" would have to discuss the entire history of the world. Many LibraryThing users, however, following

the pattern discussed by Munk and Mørk as well as Golder and Huberman, assign these broader subject terms to works that discuss a specific discipline in science or a specific time period of history.²⁰

Conversely, LibraryThing users often add terms that are more specific in nature than the subject headings supplied by catalogers. Eleven of the forty-five records (24.4 percent) examined in LibraryThing contained narrower or more specific terms than the librarian-supplied subject headings. These specific terms frequently described books that were general in nature or subject matter, such as comprehensive histories of Africa. For these books, the library catalogers typically followed standard practice and only assigned a broad heading appropriate to the overall content of the book, such as "Africa—History." LibraryThing's users, however, added terms for more specific concepts, such as "slavery," "colonialism," and "exploration." The records for Stephen J. Hawking's popular *A Brief History*

of *Time* also show this pattern. The record in the library catalog contains a single and very general subject heading, “Cosmology.” LibraryThing users, however, supplied several more specific tags, such as “physics,” “astrophysics,” “big bang,” and “black holes.”

One specific example shows all of these differences between LibraryThing tags and library-supplied subject headings. The WorldCat bibliographic record (see figure 3) for the book *A Savage War of Peace: Algeria, 1954–1962* has one LCSH, “Algeria—History—Revolution, 1954–1962.” In LibraryThing, following the expected pattern, users have assigned many more tags to the book (see figure 4). Some tags express the same concept as the LCSH but in different words, like “Algerian history,” “Algerian Revolution,” and “Algerian War.” In addition, many LibraryThing users have added tags that describe much broader concepts or that refer to broader geographic areas, like “20th Century,” “Africa,” “Middle East,” “North Africa,” “Military history,” and “war.” Other LibraryThing tags also bring out more specific concepts that may not relate to the content of the book as a whole but that some users found important, such as “colonialism,” “guerrilla,” “counterinsurgency,” and “torture.” Interestingly, in this LibraryThing record, more users assigned the tags “France” and “French history” than the tags “Algeria” and “Algerian history,” whereas the subject heading that relates to the particular war described in this book does not even mention France.

Wetterstrom’s results in the study comparing user-assigned tags and LCSH differ from results reported here.²¹ As mentioned above, Wetterstrom asked twenty people to contribute tags to a small collection of books. Wetterstrom’s study group assigned significantly fewer tags than found in this study’s LibraryThing records. Wetterstrom found an average of 24.4 user tags compared to the average of 42.78 in the LibraryThing sample. Wetterstrom’s results also differ from those seen in the present comparison of tags and LCSH. In his study, 75.47 percent of tags did not match LCSH, fewer matches than seen here. Wetterstrom’s study also found a preponderance of broader and narrower terms in the tags (14.61 percent and 19.62 percent of overall tags, respectively), although he only counted the overall number of these broader and narrower terms and did not specify if terms of these types were assigned to every book. In Wetterstrom’s study, narrower terms appeared more frequently than broader terms, which is the opposite of the results found in this study. The difference in findings may be because of two reasons. First, Wetterstrom’s study group was consciously creating tags for a research project and not for describing books in their own personal collections and which they wanted to be able to retrieve. In other words, the LibraryThing users had a personal investment in the creation of the tags and were not part of a research study group. Second, Wetterstrom’s study group (twenty individuals) was much smaller than the universe of people contributing tags

to LibraryThing, who could view tags added by others and who, in total, created significantly more tags.

Library of Congress Subject Headings

Examining the LCSH assigned to the forty-five books is instructive. Twenty-five of the WorldCat records (55.6 percent) contained LCSH expressing concepts that LibraryThing’s tags did not. These differences fall into several categories. First, LCSH in these records often refer to classes of persons, while the user tags generally only indicate abstract concepts. For example, books about Mexican immigrants have LCSH such as “immigrants” and “Mexican Americans,” nouns that refer to the groups of people. The user tags in LibraryThing for these same books, on the other hand, are “immigration” or “Mexican American,” “Chicano,” and “Latino,” that is, either the abstract concept or adjectives rather than nouns.

LCSH contains a collection of phrases known as “free-floating subdivisions.” Catalogers can append these free-floating subdivisions, with certain restrictions, to topical headings that already exist in LCSH and thereby highlight special aspects of the topical heading. Nothing like these free-floating subdivisions appears in LibraryThing’s user tags, and some records for the forty-five books considered here demonstrate the usefulness of these subheadings. Free-floating subdivisions are standardized phrases and are often not expressed in natural language. A common example of a free-floating subdivision is the term “Social conditions,” which catalogers can add to the names of places or to a phrase denoting a class of persons. The LC’s *Subject Cataloging Manual* defines this term in the following manner: “Use the subdivision for works discussing the social history or sociology of a place, ethnic group, or class of persons, including such subtopics of sociology as social problems, stability, change, interaction, adjustment, structure, social institutions, etc.”²² The term “social conditions” does not appear as a user tag, and one can see LibraryThing’s users struggling to come up with a good way to convey what the subdivision “Social conditions” expresses. Tags like “social problems,” “social history,” and “sociology” appear, but are not usually the bold-faced, more popular terms, so LibraryThing users do not seem to have a clear consensus on how to express this concept.

The tags in LibraryThing also fail to show any consensus with the expression of historical time periods, whereas in LCSH chronological divisions are established for all countries and regions. These chronological subdivisions vary from country to country and, ideally, conform to the important historical divisions within each country’s history. The library bibliographic record for *A Savage War of Peace*, as shown above, has the subject heading “Algeria—History—Revolution, 1954–1962.” The LibraryThing records for this book had very few chronological tags, and the few tags in

the record relating to time periods (“1950s,” “1960s,” “20th century,” “pre-1983,” and “post-war”) lacked specificity and consensus. The chronological subdivisions in the LCSH thesaurus do require time and effort on the part of catalogers to establish as well as to apply correctly. As the LibraryThing tags show, users do not necessarily think about historical events in neat chronological packages. The chronological subdivisions, however, like the free-floating subdivisions, do serve the purpose of bringing together materials about a given subject or, in this case, a given time period.

The LCSH system, like all thesauri and taxonomies, controls synonyms and also has an elaborate set of rules for how new subject headings are established. These rules include such basic grammatical guidelines as which types of nouns should be in the plural in subject headings and which in the singular. A subject heading that appeared frequently in this study, for example, is “Mexican Americans,” for which the LCSH thesaurus lists two nonpreferred terms, “Chicanos” and “Hispanos.” Within this system, catalogers know which term to add to a bibliographic record, and users, once they know which is the preferred term, theoretically can find all the materials under that topic. User tags in LibraryThing, however, like the tags and folksonomies studied by the authors cited above, neither control synonyms nor follow specific grammatical rules. As a result, various terms that mean more or less the same thing but are expressed differently (“Mexico,” “Mexican,” “Mexican American,” “Chicano,” and “Latino”) appear in the same LibraryThing record. This situation occurred in each of the forty-five sample records. In each record several different subject terms, or terms in different grammatical forms, expressed the same concept. Other examples include “current affairs” and “current events”; “diet,” “eating,” and “nutrition”; “Jewish,” “Jews,” and “Judaism”; “economics” and “economy”; “mountain climbing” and “mountaineering”; “decision-making” and “decisions”; and “thinking” and “thought.” Both noun and adjectival forms were used for geographic descriptors, for example, “Africa” and “African”; “Mexico” and “Mexican”; “Iran” and “Iranian.” Librarians who defend LCSH, including Mann at the LC, base their defense on the ability of subject headings to control synonyms and specify the correct grammatical form to use, which helps to collocate similar materials.²³ With all the variations possible in LibraryThing, users searching for a specific topic cannot be sure that they have found all the relevant books because they cannot necessarily predict the terms or grammatical forms that other readers have used. LibraryThing has recently added a new feature, in which users can make two tags equivalent to each other, to try to control synonyms. When that happens, only the more popular tag displays.²⁴ This interesting innovation can help reduce the redundancies in the tag cloud; however, combining tags only works when they are identical in meaning and use.²⁵ Tags like “economics” and “economy,” for example, cannot be combined. Controlled vocabularies

like LCSH, in addition to choosing preferred terms from synonyms, can also provide scope notes on how to choose between two similar but distinct terms.

Controlled vocabularies, then, help catalogers choose the appropriate subject headings to use. LibraryThing’s users, however, have an advantage over catalogers when they assign tags—they have probably read the book before tagging it on the website. Catalogers do not have the time to read an entire book before assigning subject headings and, therefore, base their subject analysis of books on words in the title, publishers’ blurbs, the preface of the book, and chapter titles. Catalogers can also be constrained by the fact that they are trying to assign subject headings that will be meaningful for a large group of unknown and potentially diverse end users, and they may not know what subjects in a book will be most important to all potential readers. In LibraryThing, the cataloger and the end user are the same.

User tags, as Spiteri has pointed out, also can adapt better and more quickly to changing terminologies and to new fields of study than LCSH or any controlled vocabulary can.²⁶ The LC has to approve new terms added to LCSH, and catalogers have to do a certain amount of research before proposing a new topical heading.²⁷ Because of this, new headings take time to appear in LCSH. In addition, LCSH are formulated to avoid polemical topics and maintain an objective stance toward the material, which often has the reverse effect of indicating a subtle bias. In the example of the book *A Savage War of Peace*, the only LCSH in the bibliographic record is “Algeria—History—Revolution, 1954–1962.” This subject heading does not explicitly mention France and its involvement in the war, while conversely more of LibraryThing’s tags for this book cited France than Algeria. The library-supplied subject heading, then, subtly erases the anticolonial nature of the war. The political implications and biases of the language used in LCSH have long interested researchers and have also inspired projects such as Berman’s alternative subject headings.²⁸ The present study is more concerned, however, with the issue of subject access to materials. A solitary subject heading like “Algeria—History—Revolution, 1954–1962” potentially hinders access to the book for readers because, according to LibraryThing tags, this book interests many users because of what it says about French history. Users browsing or searching for subjects in the library catalog for books about French military history will not come across *A Savage War of Peace*.

The subject headings in library catalog records also can suffer from what can only be called bad cataloging. Several of the books considered here have very inadequate subject headings in the library record. The book *Fast Food Nation* by Eric Schlosser provides an especially egregious example. Because this is a widely read and widely discussed book, many people, even if they have not read it, know that it is an indictment of the fast food industry that discusses the environmental, economic, and nutritional aspects of fast food

restaurants. In the library catalog, though, this book's record only has one subject heading: "Cookery, American." Even leaving aside the unnatural language in this subject heading (one example of the LC's reluctance to change headings once established), this subject heading does not accurately describe the contents of the book. It makes the book sound like a cookbook and not a social and political look at a specific part of the American food industry.

In LibraryThing, since more than one person can add tags to a record, the overall tag cloud for the book can correct a single person's errors or questionable judgment. This does point to a potential disadvantage. If only a few users have added tags, then the aggregate of tags assigned may not provide the most accurate or helpful subject analysis of the book. A few of the records under consideration here had tags supplied by only six or seven readers, and in those instances the tag clouds were not comprehensive and the tags did not provide a good description of the contents of the books. The better and more complete records in LibraryThing usually belong to more popular books, which does have certain implications for user tags in library catalogs. A large portion of the collections of university and research libraries consists of rare and little-known materials, and these are often the most useful and valuable items these libraries own. Libraries cannot rely on their users to supply the subject access for rare materials if no users, or even just a very few, have read the books. Also, the university community consists of professors and students doing higher-level research who often will need to find every book on a certain topic. To serve these scholars, libraries need to provide consistent and comprehensive subject access, and also need to use some means of controlled vocabulary that will bring together all similar materials.

Conclusion

A comparison of LibraryThing's user tags and LCSH suggests that while user tags can enhance subject access to library collections, they cannot replace the valuable functions of a controlled vocabulary like LCSH. Also, one must consider that different libraries serve different populations, and user tags will inherently be more appropriate for different libraries and types of users. Public libraries, for example, would probably benefit more readily from user tags, since their collections are often primarily popular materials. Popular books in LibraryThing tend to have tags supplied by more users, and therefore the records for these books tend to have more accurate and comprehensive tags.

If libraries do allow users to contribute tags to their catalogs, they will need to figure out how to deal with some of the inherent problems encountered in folksonomies, namely, the abundance of potentially unhelpful personal terms and the lack of control for synonyms and different

grammatical forms of words. One possibility for controlling synonyms would be to run folksonomies against automatic indexing software, but libraries will need to study when and where such an action could take place within a library's workflow and whether the results would justify taking this extra step. Many of the next-generation OPACs offer tag cloud displays that can improve the usefulness of user tags by presenting them in a way that highlights relevant terms and indicates relationships between terms.

Looking closely at the user tags in LibraryThing can also provide information on how users think about books and their subjects, and this can help improve library-supplied subject analysis, including ways in which LCSH can be improved. When library catalogs consisted of typed index cards, a conservative attitude toward changing subject headings made sense because all the cards had to be removed from the card catalog drawer and retyped. In a digital environment, however, updating bibliographic records is significantly easier, so subject headings using archaic language, such as "Cookery," do not need to be retained simply because that is how they were established. If the unnatural language in subject headings impedes access, then the headings should be updated.

The comparison of LibraryThing user tags with LCSH also shows that library catalogs do not take full advantage of all the elements already present in the subject headings system, since catalogs generally only provide an alphabetical display of subject headings. This study has shown that users assign tags that range from general to specific, whereas the subject headings assigned to bibliographic records do not cover the entire spectrum. LCSH, like most thesauri, has a hierarchical structure, with broader, narrower, and related terms indicated for most headings. This hierarchy, however, is not readily visible to most users. Libraries should consider redesigning the public display of catalogs to allow users better access to the different levels of specificity within the LC thesaurus. Some researchers have proposed enhancements to the public displays of subject headings, such as a faceted display, that would take greater advantage of the syndetic structure of LCSH.²⁹ The LC Working Group, in addition to recommending that libraries allow user tags into their catalogs, also suggests improving LCSH by allowing for these hierarchical or faceted displays.³⁰

User tags by themselves cannot provide the best subject access to the materials in library collections, but they can help point libraries in the right direction. An examination of user tags also points out the limitations of how libraries currently provide subject access to their collections. The next step, both in practical terms and as further areas of research, will be further experimentation with the inclusion of user tags into library bibliographic records and OPAC displays. LibraryThing allows libraries to display its tag clouds as part of their bibliographic records, although this display is static and does not permit users to add tags to records within the

library's catalog. Going further, the next generation of Web interfaces for catalogs is including tag clouds as part of its display and discovery options. As libraries adopt these new catalog interfaces, they will need to explore the catalogs' new discovery tools, including user tags, to see if subject access to materials is improved. The comparison of LibraryThing's user tags with LCSH shows that both types of subject access have strengths and weaknesses and suggests that libraries can best serve their users by combining different types of subject access. A combination of both types, that is, user tags to enhance discovery and controlled vocabularies to collocate like materials, may well provide the best subject access to the materials in library collections.

The present study, because it focused on a small set of books and because the books chosen were more popular than academic in nature, suggests further research could be undertaken. User tags in LibraryThing for books from special collections and from specialized academic disciplines could be studied to see if folksonomies can provide useful access for less popular materials. Tags for works of fiction, including genre fiction, also are of interest because currently LCSH are not consistently assigned to belles lettres works. Comparing tagging to cataloging in a multilingual environment or to a set of materials not in English could also be useful because controlled vocabularies like LCSH are very good at bringing together materials in different languages. Also, LibraryThing tags could be compared to a study like Wetterstrom's to see if users assign tags differently in a more controlled context. In addition, researchers could study the next generation of OPACs, which incorporate tagging and tag cloud displays, to see if these innovations help enhance discovery.

References

1. Karen Antell and Jie Huang, "Subject Searching Success: Transaction Logs, Patron Perceptions, and Implications for Library Instruction," *Reference & User Services Quarterly* 48, no. 1 (2008): 68–76.
2. Library of Congress Working Group on the Future of Bibliographic Control, *On the Record: Report of the Library of Congress Working Group on the Future of Bibliographic Control* (2008), www.loc.gov/bibliographic-future/news/lcwg-ontherecord-jan08-final.pdf (accessed Feb. 23, 2009).
3. *Ibid.*, 34.
4. Louise F. Spiteri, "The Use of Folksonomies in Public Library Catalogues," *Serials Librarian* 51, no. 2 (2006): 75–89.
5. Darlene Fichter, "Intranet Applications for Tagging and Folksonomies," *Online* 30, no. 3 (2006): 43–46.
6. *Ibid.*, 44.
7. Louise F. Spiteri, "Structure and Form of Folksonomy Tags: The Road to the Public Library Catalogue," *Webology* 4, no. 2 (2007), www.webology.ir/2007/v4n2/a41.html (accessed Oct. 1, 2008).
8. Scott A. Golder and Bernardo A. Huberman, "Usage Patterns of Collaborative Tagging Systems," *Journal of Information Science* 32, no. 2 (2006): 198–208.
9. Timme Bisgaard Munk and Kristian Mørk, "Folksonomies, Tagging Communities, and Tagging Strategies: An Empirical Study," *Knowledge Organization* 34, no. 3 (2007): 115–27.
10. *Ibid.*, 116.
11. Golder and Huberman, "Usage Patterns of Collaborative Tagging Systems."
12. Timme Bisgaard Munk and Kristian Mørk, "Folksonomy, the Power Law and the Significance of the Least Effort," *Knowledge Organization* 34, no. 1 (2007): 16–33.
13. Pauline Rafferty and Rob Hilderley, "Flickr and Democratic Indexing: Dialogic Approaches to Indexing," *Aslib Proceedings* 59, no. 4/5 (2007): 397–410; John R. Clark, "The Internet Connection: Web 2.0, Flickr and Endless Possibilities," *Behavioral & Social Sciences Librarian* 27, no. 1 (2008): 62–64.
14. Mikael Wetterstrom, "The Complementarity of Tags and LCSH: A Tagging Experiment and Investigation into Added Value in a New Zealand Library Context," *The New Zealand Library & Information Management Journal, Ng Prongo* 50, no. 4 (2008): 296–310.
15. LibraryThing, Tagging, www.librarything.com/wiki/index.php/Tagging (accessed Oct. 1, 2008).
16. LibraryThing, LibraryThing Concepts, www.librarything.com/concepts#what (accessed Oct. 1, 2008).
17. Golder and Huberman, "Usage Patterns of Collaborative Tagging Systems."
18. Emma Tonkin et al., "Collaborative and Social Tagging Networks," *Ariadne* 54 (2008), www.ariadne.ac.uk/issue54 (accessed Feb. 23, 2009).
19. Library of Congress, Cataloging Policy and Support Office, *Subject Cataloging Manual, Subject Headings* (Washington, D.C.: Library of Congress, 1996): H180.
20. Munk and Mørk, "Folksonomies, Tagging Communities, and Tagging Strategies": Munk and Mørk, "Folksonomy, the Power Law and the Significance of the Least Effort"; Golder and Huberman, "Usage Patterns of Collaborative Tagging Systems."
21. Wetterstrom, "The Complementarity of Tags and LCSH."
22. Library of Congress, Cataloging Policy and Support Office, *Subject Cataloging Manual, Subject Headings*, H2055.
23. Thomas Mann, "On the Record" but Off the Track: A Review of the Report of the Library of Congress Working Group on the Future of Bibliographic Control, with a Further Examination of Library of Congress Cataloging Tendencies (Washington, D.C.: AFSCME 2910, 2008), www.guild2910.org/WorkingGrpResponse2008.pdf (accessed Oct. 1, 2008); Thomas Mann, "The Peloponnesian War and the Future of Reference, Cataloging, and Scholarship in Research Libraries," *Journal of Library Metadata* 8, no. 1 (2008): 53–100.
24. Gene Smith, "Tagging: Emerging Trends," *Bulletin of the American Society for Information Science and Technology* 34, no. 6 (2008): 14–17.
25. LibraryThing, Tag Combining, www.librarything.com/wiki/index.php/Tag_combining (accessed Feb. 27, 2009).
26. Spiteri, "The Use of Folksonomies in Public Library Catalogues."
27. Library of Congress, Cataloging Policy and Support Office, *Subject Cataloging Manual, Subject Headings*, H202.
28. M. M. Moss, "Politically Incorrect Descriptors," *Technicalities* 17, no. 3 (1997): 12–14; Hope A. Olson, *The Power to Name: Locating the Limits of Subject Representation in Libraries* (Dordrecht, the Netherlands: Kluwer, 2002); Sanford Berman,

- Prejudices and Antipathies: A Tract on the LC Subject Heads Concerning People* (Metuchen, N.J.: Scarecrow, 1971).
29. James D. Anderson and Melissa A. Hoffman, "A Fully Faceted Syntax for Library of Congress Subject Headings," *Cataloging and Classification Quarterly* 43, no. 1 (2006): 7–38.
30. Library of Congress Working Group on the Future of Bibliographic Control, *On the Record*, 35.

Appendix. List of Books Studied

- Anzaldúa, Gloria. *Borderlands: The New Mestiza; La Frontera*. San Francisco: Spinsters/Aunt Lute, 1987.
- Atkinson, Rick. *An Army at Dawn: The War in North Africa, 1942–1943*. New York: Holt, 2002.
- Bowden, Mark. *Black Hawk Down: A Story of Modern War*. New York: Atlantic Monthly, 1999.
- Boyle, T. Coraghessan. *The Tortilla Curtain*. New York: Viking, 1995.
- Bryson, Bill. *A Short History of Nearly Everything*. New York: Broadway, 2003.
- Conover, Ted. *Coyotes: A Journey through the Secret World of America's Illegal Aliens*. New York: Vintage, 1987.
- Davidson, Basil and the editors of Time-Life Books. *African Kingdoms*. New York: Time-Life, 1971.
- Diamond, Jared. *Guns, Germs, and Steel: The Fates of Human Societies*. New York: Norton, 2005.
- Ehrenreich, Barbara. *Nickel and Dimed: On (Not) Getting by in America*. New York: Metropolitan, 2001.
- Fanon, Frantz. *The Wretched of the Earth*. Trans. Constance Farrington. New York: Grove, 1968.
- Frank, Anne. *The Diary of a Young Girl*. Trans. B. M. Mooyaart-Doubleday. Garden City, N.Y.: Doubleday, 1952.
- Gladwell, Malcolm. *Blink: The Power of Thinking without Thinking*. New York: Little, Brown, 2005.
- Gonzalez, Juan. *Harvest of Empire: A History of Latinos in America*. New York: Viking, 2000.
- Gourevitch, Philip. *We Wish to Inform You that Tomorrow We Will Be Killed with Our Families: Stories from Rwanda*. New York: Farrar, 1998.
- Gutierrez, David G., ed. *Between Two Worlds: Mexican Immigrants in the United States*. Wilmington, Del.: Scholarly Resources, 1996.
- Haley, Alex. *Roots*. Garden City, N.Y.: Doubleday, 1976.
- Hanson, Victor Davis. *Mexifornia: A State of Becoming*. San Francisco: Encounter, 2003.
- Hawking, Stephen W. *A Brief History of Time: From the Big Bang to Black Holes*. Toronto: Bantam, 1988.
- Hochschild, Adam. *King Leopold's Ghost: A Story of Greed, Terror, and Heroism in Colonial Africa*. Boston: Houghton Mifflin, 1998.
- Horne, Alistair. *A Savage War of Peace: Algeria, 1954–1962*. New York: Penguin, 1987.
- Jiménez, Francisco. *Breaking Through*. Boston: Houghton Mifflin, 2001.
- Krakauer, Jon. *Into Thin Air: A Personal Account of the Mount Everest Disaster*. New York: Villard, 1997.
- Krakauer, Jon. *Under the Banner of Heaven: A Story of Violent Faith*. New York: Doubleday, 2003.
- Larson, Erik. *The Devil in the White City: Murder, Magic, and Madness at the Fair that Changed America*. New York: Crown, 2003.
- Levitt, Steven D., and Stephen J. Dubner. *Freakonomics: A Rogue Economist Explores the Hidden Side of Everything*. New York: William Morrow, 2005.
- Louie, Miriam Ching Yoon. *Sweatshop Warriors: Immigrant Women Workers Take on the Global Factory*. Cambridge, Mass.: South End, 2001.
- Mandela, Nelson. *Long Walk to Freedom: The Autobiography of Nelson Mandela*. Boston: Little, Brown, 1994.
- Martínez, Rubén. *Crossing Over: A Mexican Family on the Migrant Trail*. New York: Metropolitan, 2001.
- Meredith, Martin. *The Fate of Africa: From the Hopes of Freedom to the Heart of Despair; A History of Fifty Years of Independence*. New York: Public Affairs, 2005.
- Moorehead, Alan. *The Blue Nile*. New York: Harper & Row, 1962.
- Moorehead, Alan. *The White Nile*. New York: Harper, 1960.
- Morris, Donald R. *The Washing of the Spears: A History of the Rise of the Zulu Nation and Shaka and Its Fall in the Zulu War of 1879*. New York: Simon & Schuster, 1965.
- Nafisi, Azar. *Reading Lolita in Tehran: A Memoir in Books*. New York: Random House, 2003.
- Nazario, Sonia. *Enrique's Journey*. New York: Random House, 2006.
- Pakenham, Thomas. *The Boer War*. New York: Random House, 1979.
- Pakenham, Thomas. *The Scramble for Africa, 1876–1912*. New York: Random House, 1991.
- Roach, Mary. *Stiff: The Curious Lives of Human Cadavers*. New York: Norton, 2003.
- Ryan, Pam Muñoz. *Esperanza Rising*. New York: Scholastic, 2000.
- Schlosser, Eric. *Fast Food Nation: The Dark Side of the All-American Meat*. Boston: Houghton Mifflin, 2001.
- Thompson, Gabriel. *There's No José Here: Following the Hidden Lives of Mexican Immigrants*. New York: Nation, 2007.
- Truss, Lynne. *Eats, Shoots & Leaves: The Zero Tolerance Approach to Punctuation*. New York: Gotham, 2004.
- Urrea, Luis Alberto. *Across the Wire: Life and Hard Times on the Mexican Border*. New York: Anchor, 1993.
- Urrea, Luis Alberto. *The Devil's Highway: A True Story*. Boston: Little, Brown, 2004.
- Ward, Elijah. *Narcocorrido: A Journey into the Music of Drugs, Guns, and Guerrillas*. New York: Rayo, 2001.
- Winchester, Simon. *The Professor and the Madman: A Tale of Murder, Insanity, and the Making of the "Oxford English Dictionary"*. New York: HarperCollins, 1998.