User Tagging Behaviors in an OPAC

An Analysis of Seven Years of I-Share User Tags

Brinna Michael and Myung-Ja Han

User tagging services are underused in cultural heritage institutions despite their availability for over a decade. This study considers seven years of user tags from university and public institutions by comparing tagging service usage between institution types and qualitatively analyzing a selection of tags from the University of Illinois. Researchers found that overall, few users tag items in online catalogs, but those tags that are being created are largely descriptive in nature, indicating the potential to improve discoverability for underdescribed materials, e.g., lack of subject headings. With improved education on their use and purpose, tagging and annotation services can become important resources for cultural heritage institutions.

Discoverability access service is at the heart of the library's daily functions and depends largely on discovery systems, including the online access catalog (OPAC) and metadata, notably MARC records. As technologies advance, new and innovative opportunities arise to enhance access and discovery layers, and libraries have diligently experimented with them to adapt to some of these changes. One example is the user tagging service, a function that stemmed from the phenomenon of social tagging on the open web, often referred to as Web 2.0. User tagging has generated excitement and controversy in technical services because of the question: what role do uncontrolled user tags play in improving discovery and access in comparison to and in conjunction with the existing authority control of cataloging standards and practices?

This study explored user behavior when given the opportunity to tag within an OPAC environment and examined the purpose and reality of user tagging as a complementary service to traditional cataloging. Specifically, this study intended to capture and assess aspects of the context under which users are tagging materials, including categorizing tags based on their relationship to existing descriptive metadata and contextual relevance. To do this, researchers worked with the Consortium of Academic and Research Libraries in Illinois (CARLI) to gather bibliographic records and associated tags from the I-Share integrated library system and its VuFind discovery layer. First, the data were assessed as a whole to determine the distribution and frequency of user tagging across institution types. Next, a sample of the data was taken to classify and analyze tags in their context within the OPAC.

Brinna Michael (bamichael@emory.edu) is Cataloging and Metadata Librarian, Pitts Theology Library, Emory University; **Myung-Ja Han** (mhan3@illinois.edu) is Head, Acquisitions and Cataloging Services, University of Illinois at Urbana-Champaign.

Literature Review

In his paper, "Tagging for Libraries: A Review of the Effectiveness of Tagging Systems for Library Catalogs," Gerolimos outlined the emergence of trends within the study of tagging in information sciences literature. He addressed the increase of interest in tagging that began in the mid-2000s following the success of social networking sites like Facebook and Twitter.[1] He tracked the shifts in research trends in the late 2000s and early 2010s towards implementation of tagging services within libraries and on websites dedicated to more traditional library materials, like Goodreads and LibraryThing.[2] During this period, there was an emphasis on the comparison between user generated tags and controlled vocabularies, primarily the Library of Congress Subject Headings (LCSH), and divided perspectives on the validity and usefulness of the folksonomies for search and discovery.[3]

As Gerolimos's review revealed, librarians and other information professionals were concerned with the nature of tags as an uncontrolled vocabulary, though many recognized the potential benefits, including a more inclusive vocabulary of description, facilitating serendipitous discovery, and the potential to alleviate costs when the implementation of a controlled vocabulary is not viable.[4] He concluded that research on the use of tags in the library catalog should reach beyond "determining the quality of user tags compared to subject headings," and expand to answer broader questions:[5]

> How did the tag system manage to transfer that feeling of "importance" in creating online content and describing resources to its users...? To what extent is the effort of tag assignment to document records based on real-time need to augment the search capabilities of OPACs? At what level are users infused with the willingness to provide keywords to enhance . . . the search/research options of other users with the use of tags? And how likely is it that the subsequent user will benefit from the keywords chosen by the one before him?[6]

Since Gerolimos's review, researchers have expanded the breadth of their inquiry into tagging and the behaviors surrounding the practice. Syn and Spring addressed methods for determining the potential of user generated tags to classify a collection based on metrics intended to determine user agreement and remove terms that are too broad or narrow.[7] Joorabchi, English, and Mahdi investigated the feasibility of integrating tags and linked data methods to improve issues of inconsistency within such uncontrolled, but valuable, vocabularies. Still other researchers have studied influences on user tagging behaviors in a variety of environments, focusing on the motivations behind the act of tagging itself.[8]

This study's scope was to expand upon such research, interrogating and applying observations on user tagging behaviors broadly. In analyzing these behaviors, researchers looked back and expanded on previous investigations into the relative value of and usability of user tags as a unique descriptive resource alongside traditional cataloging, addressing several of the questions Gerolimos proposed. This study focused on the tagging behaviors of users in academic library OPACs, and considers the context within which tags are made, the type of tag, and the implications of user tagging trends. As a result, the researchers designed this study to address the following questions:

- To what degree are users adding tags in an OPAC if the system allows such functionality?
- What types of tags are being added and in what context?

Additionally, the researchers sought to explore how this study might inform current discussion surrounding the following questions:

- Can libraries utilize user added tags to improve discovery and access services?
- Are tagging services still valid and useful in the age of linked open data?

Method

For the purposes of this study, CARLI provided researchers with data in the form of a tab delimited file, listing as one unit the bibliographic record number and prefix indicating the holding institution, the number of users who had added tags, the total number of tags added, and a list of all tags added to the record. The data was drawn from eighty-nine institutions participating in I-Share, the collective integrated library system and shared OPAC offered by CARLI, and reflected all tags created from the service's implementation of the VuFind discovery layer from June 2010 to March 2017, when the data were collected. Due to the nature of the data, researchers identified four data types: institution, bibliographic record, number of users who added a tag(s) to a record, and the tag(s) added. By defining these data types, researchers were able to both examine the individual types and the relationships between each type.

Having arranged the data in this manner, the researchers designed a two-part approach to the data analysis. First, researchers grouped the data based on institution type using the Carnegie Classification of Institutions' Basic Classification guidelines to conduct a quantitative analysis of all data types.[9] The Carnegie Classification of Institutions was selected for its consistency and accuracy as an ongoing

standard of categorization of institutions of higher education. Second, a sample set of the data was identified and the associated tags categorized based on a set of categories identified by the researchers.

For the first analysis, the data consisted of 286,805 tags, 157,215 records, and 167,095 users from eighty-nine institutions. The institutions were divided into groups based on the five Basic Classifications defined by the Carnegie Classification: Doctoral Universities, Master's Colleges and Universities, Baccalaureate Colleges, Associate's Colleges, and Special Focus/Other.[10] Within these categories, the total tags, users, and records were compiled for each individual institution, the five institution categories, and the data set as a whole (see Appendix A). These totals were used to calculate the number of tags appearing per record on average, the number of users adding tags per record on average, and the number of tags being added per user on average. These three averages were calculated for individual institutions, institutional categories, and the data set as a whole.

For the second analysis, data from the University of Illinois (U of I) was selected as a sample from the full CARLI data (see table 1). To work with this sample, researchers isolated the bibliographic record numbers for the records associated with U of I and ran a report to pull the associated MARC 245 ($a and $b), 100 ($a), 650 (all subfields), 651 (all subfields), and 655 (all subfields) data fields that represent the title, author, and subjects of each record. The resulting data set was compiled and uploaded into OpenRefine, an open source application for data cleaning and exploration. The researchers used the faceting feature to identify records that lacked values in the 650, 651, or 655 fields (i.e., any subject headings). These records were chosen for the sample and resulted in 2,605 tags, 1,237 users, and 1,207 records.

To contextualize the tags associated with U of I's sample, OpenRefine's faceting and clustering functions were used to produce a list of unique tags. In OpenRefine, the faceting function identifies each unique string value in a column and returns the number of times each string appears in the column. The clustering function can then be used to reconcile string values that are marked as similar according to an algorithm that determines "sameness" using a key collision method called fingerprinting.[11] For this process, the researchers removed extra whitespace and punctuation at the beginning and end of strings. No tags were changed in regard to case or spelling to retain as much original context as possible.

Researchers then performed a cursory overview of the resulting list of unique tags and identified common themes from which categories could be determined. Based on these observations, researchers identified seven clear categories (see table 2). All tags remaining after the initial sort were assessed against their full bibliographic record and sorted to the best of the researchers' abilities. The remaining tags following this secondary sort were grouped into a final category, Other.

Results

Institutional Classification

Of the eighty-nine institutions identified within the data set, researchers identified ten doctoral universities, twenty-five master's colleges and universities, fourteen baccalaureate

Table 1. University of Illinois at Urbana-Champaign Full and Sample Data

	Total U of I data	U of I Data without Subject Headings
Data types		
Records	21,776	1,207
Users	22,863	1,245
Tags	37,706	2,595
Unique tags	8,883	1,083
Tags per record		
Minimum	1	1
Maximum	37	14
Average	1.732	2.136

Table 2. Tag Categories

Category	Definition	Example
Content Description	Describes or addresses what the work is "about"	action, romance
Title Words	Matches a word(s) in the title of the work as it appears in the 245 field	Bhagwad Gita
Creator Name	Matches the name(s) of the work's creator(s) as they appear in the 100 field	kafka, Calvino
User Commentary	User notes, intentions, actions, and evaluations	diss, REQUEST
Course Information	Indicates a course name and/or number	ARTF101, AmLit
Object Description	Describes or addresses the physical or digital object	e-book, map
Call Number/Location	Indicates the call number or physical location of the object	L-OSF, stacks

colleges, twenty-four associate's colleges, and sixteen special focus/other. After classifying all institutions, the number of individual tags, records, and users were quantified at the institution level and then averaged within each category. These results showed that on average, institutions classified as doctoral universities had the highest record, user, and tag counts when compared to other institutions and accounted for 52 percent of all records, 63 percent of all users, and 54 percent of all tags (see figure 1).

Despite representing only 11 percent of the participating institutions, doctoral universities were responsible for the bulk of the cumulative data in all three types. This phenomenon reflected the relative sizes of these institutions when considering the number of students, staff, and faculty (users) and volumes held (records). Larger collections and a greater number of potential users increase the overall tag output. The discrepancy in size of the collection and potential user pool between institution types did not appear to affect the likelihood of users adding tags to records as evidenced by an assessment of the relationships between each data type (see figure 2).

As illustrated in figure 2, researchers calculated the average number of users adding tags per record, tags added per record, and user to tag ratio. These relationships did not show a significant variation across institution types, thereby indicating a consistency with which users across institution types applied tags to records. This trend exhibited an independence from the relative size of the potential user group or institutional collection.

Subset Determination

To analyze the tags, researchers extracted data associated with U of I. U of I was categorized as a doctoral university and had 21,776 records, 22,863 users, and 37,706 tags total. Compared to other doctoral universities, the ratios of users per record (1.05:1), tags per record (1.732:1), and tags per user (1.649:1) for U of I's data was well within the expected results.

Of the 21,776 records, researchers identified 1,207 records lacking subject headings, representing approximately 6 percent of the U of I data and 0.8 percent of the full I-Share data (see table 1). There are some brief records, and others are for literature that normally do not have subject headings. These records were extracted as a subset of the full data to be used for qualitative analysis on the basis that users would have tagged these materials under significantly less influence by the catalog records. The same quantitative analyses as the full data set was applied and compared to the rest of the U of I data.

In a comparison of the records lacking subject headings against the full U of I records, on average those without subject headings had a higher ratio of tags per record (2.136:1).

When comparing the number of tags per record, both sets showed similar trends. As shown in figure 3, an analysis of the number of tags per record for the total records from U of I showed that approximately 62.12 percent of records had only one tag, while the maximum number of tags for a single record was thirty-seven. Comparatively, when only the records lacking subject headings were analyzed, approximately 62.06 percent of the records had only one tag, while the maximum number of tags for a single record was fourteen.

Tag Categorization

After sorting tags into the previously identified eight categories, researchers analyzed the resulting groupings and found that tags fell overwhelmingly into the Content Description category (54.22 percent). The second largest category, Title Words (22.04 percent), included a number of tags that could logically have been categorized as Content Description on the basis that titles are generally considered to be descriptive of a work's contents. Researchers determined that the majority of the tags broadly described the contents of the resources (see table 3). The prevalence of descriptive tags indicated that many users have clear objectives when they added tags.

To further analyze the results of categorization, researchers extracted lists of all unique tags and their frequency of occurrence from the I-Share data, the U of I data, the full set of records without subject headings, and those tags categorized under Content Description (see Appendix B). In comparing the top thirty most frequently occurring tags, researchers recognized a variation in the specificity of the tags from the full data set and the U of I data and those from the subset and Content Description category. The tags for the I-Share records and the U of I records appeared to be more general, with some user commentary such as "to read" and initials, plus notes about the item's intended use ("research" or "paper"). The subset and Content Description tags exhibited a greater degree of specificity, focused more on describing the genre of the resources with terms such as "Drama," "comedy," and "romance." This sharpening of specificity indicated to researchers that users' descriptive tagging behaviors became more pointed and purposeful when the subject headings in the catalog records were limited or non-existent.

Discussion

User tagging has a long history of debate among the cultural heritage community in relation to the service's potential for enhancing access and discoverability of materials. Assessment of the I-Share user tags indicated a limited use of
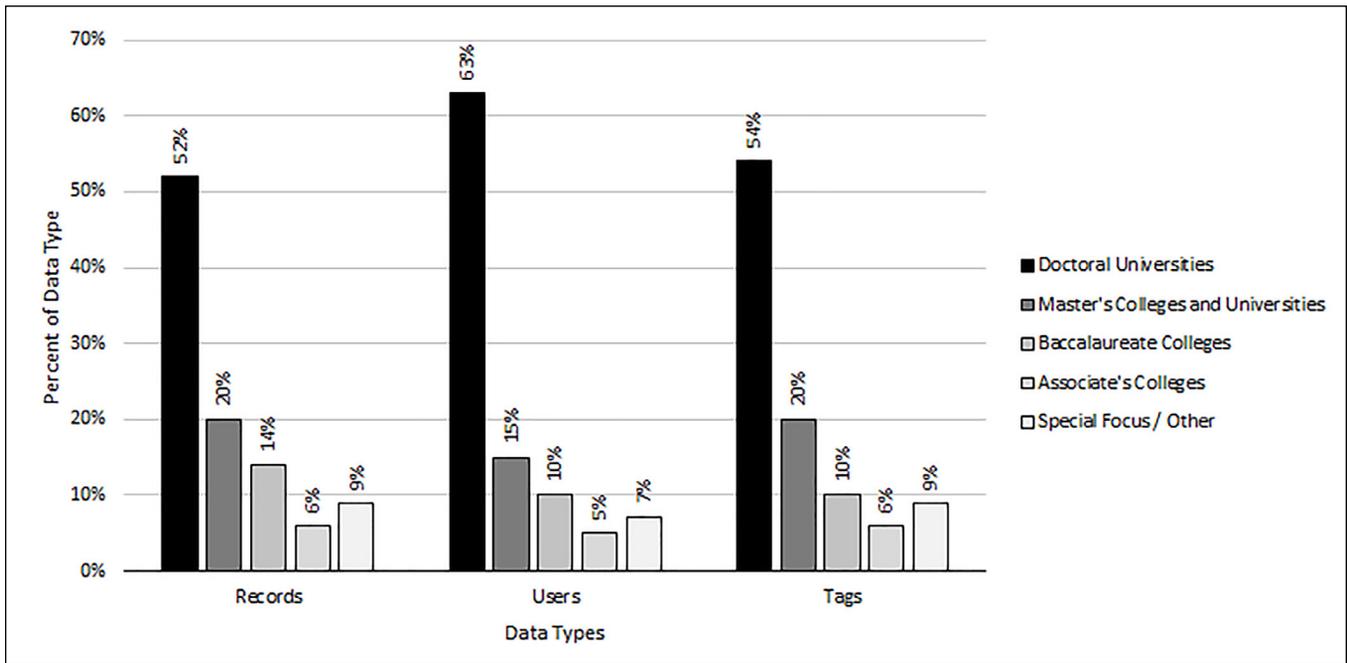
Figure 1. Percent of Cumulative Data by Institution Type
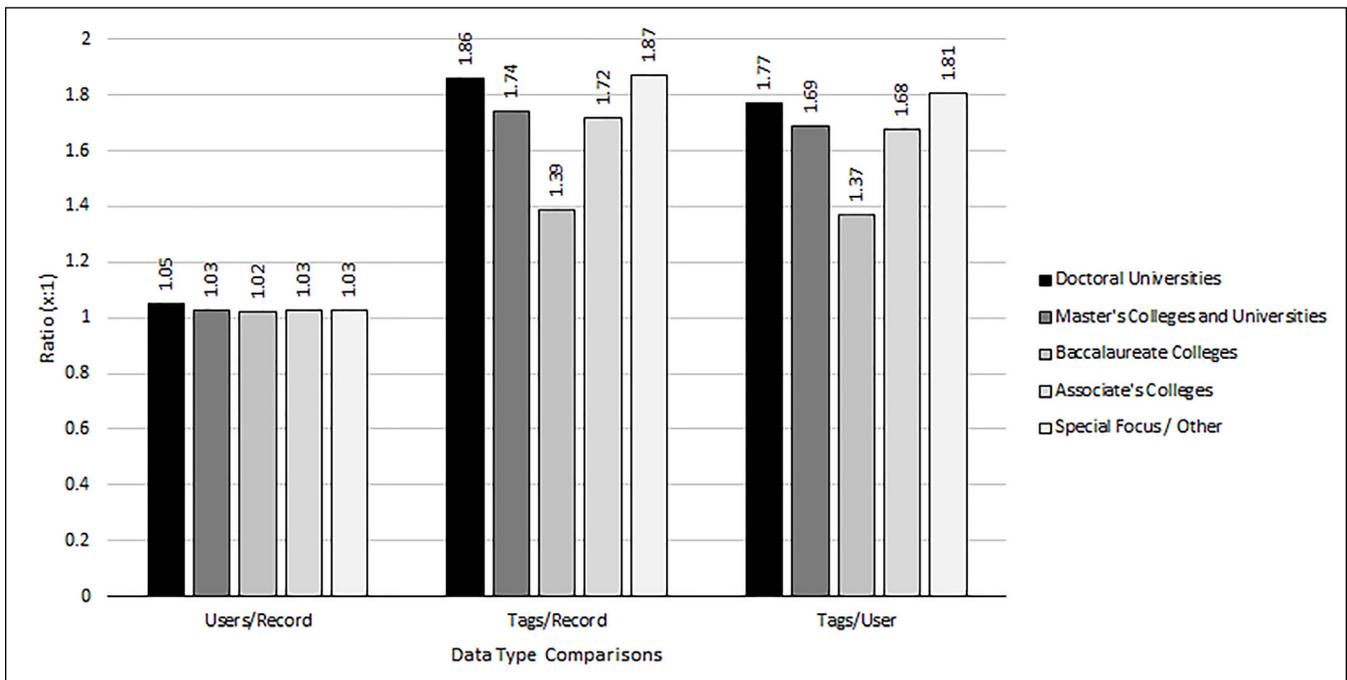
Figure 2. Relationship between Data Types

tagging services by users across academic institution types, with the likelihood of users to tag remaining relatively standard across institution types. Although 157,215 individual item records were represented in this study, this is a modest percentage of the combined holdings of the

eighty-nine participating institutions that represent a collective 14.7 million unique bibliographic records and 38.1 million item records. The reasons for such a small portion of materials being tagged could be attributed to a number of factors: lack of user awareness of tagging services, lack of

Figure 3. Frequency of Tags per Record

user education on the use of tagging services, lack of user interest in tagging services, lack of use cases on how to use user tags in cataloging or (and) discovery services, etc. Regardless, several trends emerged from the data collected via I-Share that merit discussion.

User purpose for tags appears varied but can largely be understood to fall into three behaviors: adding context to described or under-described materials, creating a personal collection for research or reference, and indicating personal perception and/or future 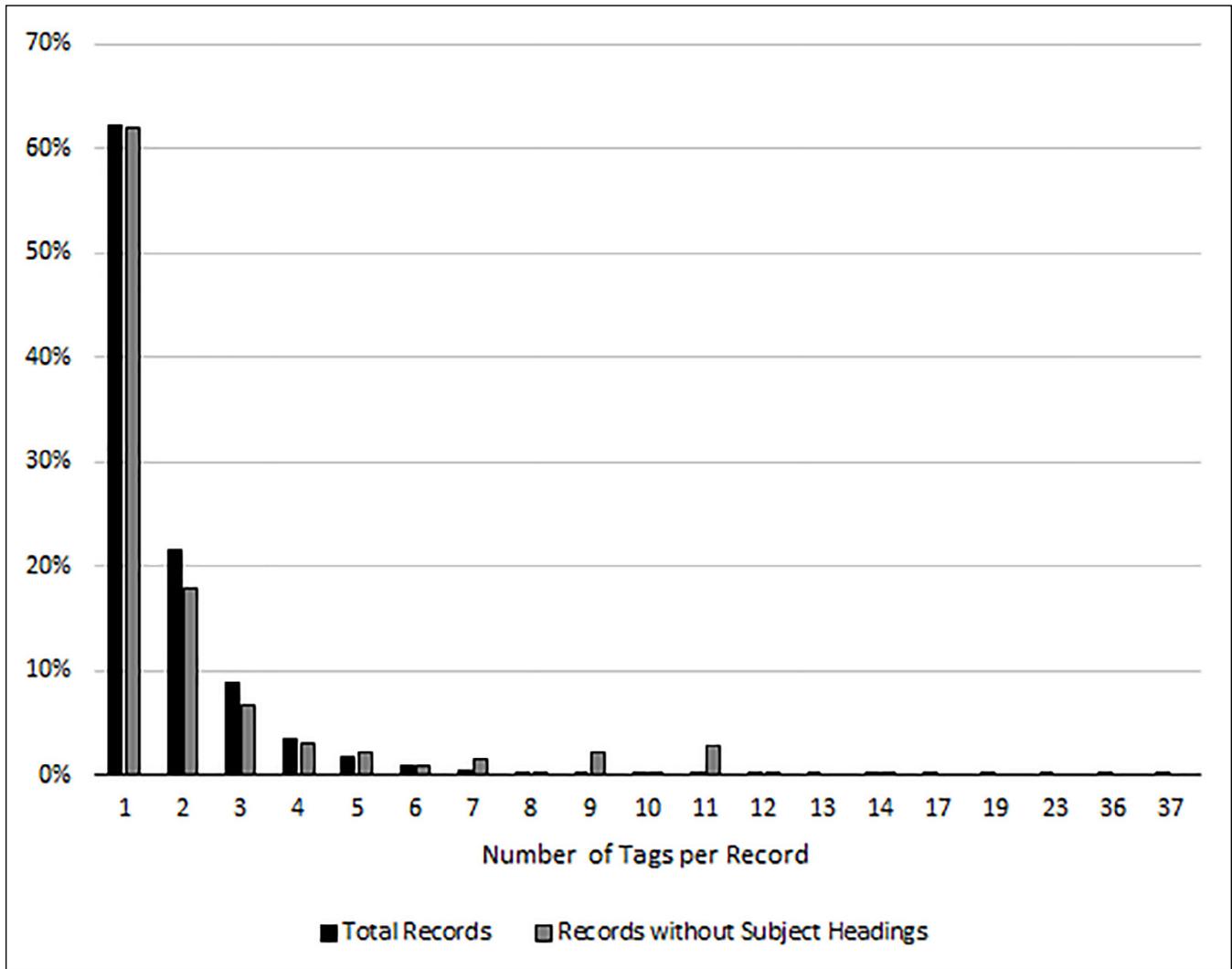intentions. The presence of tags such as "jkbnhs," which appears a total of 327 times in the full I-Share data set, indicates a behavior of collecting materials through personalized tags. Additionally, tags such as "diss" and "ARTF101" indicate a variation on this collecting behavior, grouping items based on relevance to research or coursework.

The prevalence of descriptive tags indicates a desire

to enhance the description of records both for public and personal use. Annotations have been broadly defined to include any type of marking or notation made with the purpose of indicating observations, comments, and intentions. Using this definition suggests that the behaviors of users tagging records in the OPAC is a form of annotation with limited functionality. One constraint on the functionality of VuFind's tagging service is how tags are processed and added to the catalog. To add a single word tag, users need only type the word into the designated search box. To add a phrase, users must enclose the phrase in quotes (see table 4).

The result is that some users appear to have followed the input requirements for phrases, while others did not, resulting in several individual tags, that when read together, complete a full annotative thought. These actions account for the variation in the number of tags per record and supports the observation that a lack of user education on how

Table 3. Results of Tag Categorization

Category	Tag Count	% of Total Tags
Content Description	1,407	54.22
Title Words	572	22.04
User Commentary	198	7.63
Creator Name	173	6.67
Course Information	85	3.28
Other	76	2.93
Object Description	58	2.24
Call Number/Location	26	1.00

* Note: Percentages calculated using the number of tags from the U of I Data without Subject Headings (see table 1).

Table 4. User Input Effect on Tag Output

User Input	Resulting Tag(s)
book	book
"to check out"	to check out
things I'm interested in	things, I'm, interested, in

to use tagging services plays a role not only in the perception of the nature and meaning of a tag or tags, but also in the interpretation of the relevancy of tags to both users and library staff as evidenced by the researchers' disregard for individual tags that are considered stop words in the analysis of the most frequently occurring tags, important context is lost without a reassessment of the context in which those tags exist.

Conclusion

When first introduced in the early 2000s, user tagging services were regarded as one of the direct implementations of Web 2.0 utility and welcomed by the library and cultural heritage community.[12] This study examined users' tagging behaviors in an OPAC by analyzing user tags added to the CARLI integrated library system from 2010 to 2017. Data analysis revealed that the tagging service is not used as much as anticipated, and that only a small number of CARLI records include user tags.

When examined closely, the study found that users create tags largely for descriptive purposes, although many tags indicate personal annotation when applied. This trend has led some researchers to speculate whether user tagging services is no longer desirable in the era of linked open data. However, based on this study's findings, researchers believe there are ways to improve user tagging services. They encourage libraries to explore other options that facilitate the incorporation of user tagging into the main library services.

First, the analysis revealed that users added tags for a variety of purposes, all of which could be broadly considered annotations. Recently, the W3C Annotation Group published a data model and vocabularies for the web annotation service.[13]

Second, based on the limited use of user tagging services and the generally low quality of tags, libraries should seek to improve user education on the use and purpose of tagging and/or annotating in the OPAC. Users cannot use the service to full advantage nor provide quality tags when they are not aware of the service or how to use it. Coordinated instruction opportunities with public services or library instruction departments and a readily useable web document could provide the education necessary to fully utilize tagging or annotation services.

Third, because tags are uncontrolled, there is a certain limitation on integrating tags into a library's bibliographic records. However, tags could still be used as part of the discovery services. VuFind version 4.3 includes user tags as a search options, in addition to more traditional search methods.[14] The inclusion of tags as an indexed and searchable information source may aid users in discovering items when using natural language queries that are more familiar to them than library specific controlled vocabularies, such as Library of Congress Subject Headings. Because user tags tap into users' natural language habits, they not only provide an alternate descriptive vocabulary, but also capture the unique perspectives and language of the users providing them.

While user-tagging services have been available since the early 2000s, they are underused for various reasons. As libraries and other cultural heritage institutions move towards adopting linked data and web technologies, it is time to reevaluate the service and find ways to better integrate tags, as a unique and user-reflective resource, into our discovery services to improve access to under-cataloged library materials and promote scholarly communication.

References and Notes

1. Michalis Gerolimos, "Tagging for Libraries: A Review of the Effectiveness of Tagging Systems for Library Catalogs," *Journal of Library Metadata* 13, no. 1 (2013): 37.
2. Gerolimos, "Tagging for Libraries," 38.
3. Gerolimos, "Tagging for Libraries," 39–48.
4. Gerolimos, "Tagging for Libraries," 42–43, 45–47.
5. Gerolimos, "Tagging for Libraries," 51.
6. Gerolimos, "Tagging for Libraries," 51–52.

7. Sue Yeon Syn and Michael B. Spring, "Finding Subject Terms for Classificatory Metadata From User-Generated Social Tags," *Journal of the Association for Information Science & Technology* 64, no. 5 (2013): 964–80.

8. Yi-ling Lin et al., "The Impact of Image Descriptions on User Tagging Behavior: A Study of the Nature and Functionality of Crowdsourced Tags," *Journal of the Association for information Science & Technology* 66, no. 9 (2015): 1785-1798; Youngok Choi and Sue Yeon Syn, "Characteristics of Tagging Behavior in Digitized Humanities Online Collections," *Journal of the Association for Information Science & Technology* 67, no. 5 (2016): 1089–104; Youngok Choi, "The Nature of Tags in a Knowledge Organization System of Primary Visual Resources," *Journal of Library Metadata* 17, no. 1 (2017): 37–53.

9. "Basic Classification Description," Definitions, The Carnegie Classification of Institutions, accessed September 20, 2017, http://carnegieclassifications.iu.edu /classification_descriptions/basic.php.

10. The last category, Special Focus/Other, includes a number of institutions that do not fall within the purview of the Carnegie Classification of Institutions including EBL PDA eBooks, HathiTrust Digital Library, Illinois Math and Science Academy, Illinois State Library, JKM Library Trust, and the Newberry Library.

11. Owen Stephens, "Clustering In Depth," OpenRefine, last modified May 13, 2018, https://github.com/OpenRefine /OpenRefine/wiki/Clustering-In-Depth.

12. Choi and Syn, "Characteristics of Tagging Behavior," 1089–90; Choi, "The Nature of Tags," 37–38.

13. Coralie Mercier, "Three Recommendations to Enable Annotations on the Web," W3C, last modified February 23, 2017, www.w3.org/blog/news/archives/6156.

14. Villanova University, "VuFind, Search, Discover, Share," last modified May 15, 2019, https://vufind.org/vufind/features.html.

Appendix A. I-Share Data Types by Institution and Institution Classification

	Institutions	Records	Users	Tags	User/ Records	Tag/ Records	Tag/User
Total	89	157,215	167,095	286,805	1.063	1.825	1.716
Doctoral Universities	10	48,301	50,715	89,892	1.045	1.954	1.866
Benedictine University		598	629	1,260	1.052	2.107	2.003
DePaul University		567	569	1,020	1.004	1.799	1.793
Illinois Institute Of Technology		3,760	4,274	8,587	1.137	2.284	2.009
Illinois State University		3,270	3,301	5,584	1.009	1.708	1.692
Northern Illinois University		4,559	4,637	8,443	1.017	1.852	1.821
National Louis University		956	1,014	2,191	1.061	2.292	2.161
Southern Illinois University Carbondale		3,329	3,383	5,963	1.016	1.791	1.763
Trinity International University		2,622	2,735	5,026	1.043	1.917	1.838
University Of Illinois Chicago		6,864	7,310	14,112	1.065	2.056	1.931
University Of Illinois Urbana-Champaign		21,776	22,863	37,706	1.05	1.732	1.649
Master's Colleges And Universities	25	18,893	19,439	32,927	1.015	1.784	1.758
Aurora University		508	516	932	1.016	1.835	1.806
Bradley University		836	847	1,569	1.013	1.877	1.852
Columbia College Chicago		809	812	1,391	1.004	1.719	1.713
Concordia University Chicago		115	115	204	1	1.774	1.774
Chicago State University		89	90	250	1.011	2.809	2.778
Dominican University		1,029	1,067	1,688	1.037	1.64	1.582
Eastern Illinois University		2,261	2,295	3,184	1.015	1.408	1.387
Elmhurst College		273	272	470	0.996	1.722	1.728
Greenville University		155	159	293	1.026	1.89	1.843
Governors State University		748	754	1,186	1.008	1.586	1.573
Judson University		444	450	782	1.014	1.761	1.738
Lewis University		137	138	276	1.007	2.015	2

	Institutions	Records	Users	Tags	User/ Records	Tag/ Records	Tag/User
Mckendree University		147	149	267	1.014	1.816	1.792
North Central College		506	511	911	1.01	1.8	1.783
Northeastern Illinois University		989	1,003	1,846	1.014	1.867	1.84
North Park University		2,518	2,697	3,660	1.071	1.454	1.357
Olivet Nazarene University		1,892	1,986	3,689	1.05	1.95	1.858
Quincy University		229	229	301	1	1.007	1.007
Robert Morris University		145	145	238	1	1.641	1.641
Roosevelt University		1,332	1,360	2,699	1.021	2.026	1.985
Southern Illinois University Edwardsville		2,826	2,936	5,517	1.039	1.952	1.879
Saint Xavier University		163	164	236	1.006	1.448	1.439
University Of Illinois Springfield		469	471	798	1.004	1.701	1.694
University Of St. Francis		203	203	407	1	2.005	2.005
Western Illinois University		70	70	133	1	1.9	1.9
Baccalaureate Colleges	14	12,742	12,959	17,674	1.018	1.648	1.62
Augustana College		392	395	612	1.008	1.561	1.549
Eureka College		154	158	269	1.026	1.747	1.703
Illinois College		6,482	6,503	7,110	1.003	1.097	1.093
Illinois Wesleyan University		412	413	794	1.002	1.927	1.923
Kendall College		80	81	130	1.013	1.625	1.605
Knox College		1,657	1,712	2,556	1.033	1.543	1.493
Lake Forest College		474	472	815	0.996	1.719	1.727
Lincoln College		305	305	306	1	1.003	1.003
Millikin University		828	856	1,425	1.034	1.721	1.665
MacMurray College		7	7	12	1	1.714	1.714
Monmouth College		203	205	380	1.01	1.872	1.854
Principia College		575	568	1,150	0.988	2	2.025
Trinity Christian College		288	291	490	1.01	1.701	1.684
Wheaton College		885	993	1,625	1.122	1.836	1.636
Associate's Colleges	24	5,620	5,756	9,663	1.009	1.742	1.73
Black Hawk College		9	9	21	1	2.333	2.333
College Of DuPage		340	341	505	1.003	1.485	1.481
Carl Sandburg College		14	13	25	0.929	1.786	1.923
Danville Area Community College		37	36	80	0.973	2.162	2.222
Heartland Community College		242	248	460	1.025	1.901	1.855
Illinois Central College		876	882	1,681	1.007	1.919	1.906
Illinois Eastern Community Colleges*		105	105	194	1	1.848	1.848
Illinois Valley Community College		445	461	722	1.036	1.622	1.566
Joliet Junior College		383	388	633	1.013	1.653	1.631
John Wood Community College		1	1	1	1	1	1
Kankakee Community College		19	19	27	1	1.421	1.421
Kishwaukee College		159	163	241	1.025	1.516	1.479
Lewis And Clark Community College		162	164	377	1.012	2.327	2.299
Lincoln Land Community College		192	194	364	1.01	1.896	1.876

Institutions		Records	Users	Tags	User/Records	Tag/Records	Tag/User
Morton College		21	21	32	1	1.524	1.524
Oakton Community College		678	691	1,096	1.019	1.617	1.586
Parkland College		493	496	816	1.006	1.655	1.645
Richland Community College		89	90	128	1.011	1.438	1.422
Southeastern Illinois College		1	1	1	1	1	1
South Suburban College		3	3	11	1	3.667	3.667
Sauk Valley Community College		625	701	839	1.122	1.342	1.197
Southwestern Illinois College		111	111	195	1	1.757	1.757
Triton College		77	77	100	1	1.299	1.299
(William Rainey) Harper College		538	541	1,114	1.006	2.071	2.059
Special Focus/Other	16	8,007	8,282	14,956	1.033	1.816	1.757
Adler University		438	465	853	1.062	1.947	1.834
Chicago School Of Professional Psychology		78	4	182	0.051	2.333	45.5
Catholic Theological Union		499	506	850	1.014	1.703	1.68
Northern (Baptist Theological) Seminary		207	207	517	1	2.498	2.498
University Of Saint Mary Of The Lake (Mundelein Seminary)		187	189	296	1.011	1.583	1.566
Harrington College Of Design		314	324	488	1.032	1.554	1.506
Lincoln Christian University		444	448	749	1.009	1.687	1.672
School Of The Art Institute Of Chicago		1,587	1629	3,024	1.026	1.905	1.856
Rush University		89	103	188	1.157	2.112	1.825
Southern Illinois University School Of Medicine		111	112	129	1.009	1.162	1.152
EBL PDA Ebooks		1,724	1,932	3,527	1.121	2.104	1.877
HathiTrust		1,626	1,655	2,941	1.018	1.809	1.777
Illinois Math And Science Academy		278	276	355	0.993	1.277	1.286
Illinois State Library		175	182	334	1.04	1.909	1.835
JKM Library Trust		157	157	371	1	2.363	2.363
Newberry Library		93	93	152	1	1.634	1.634

* Illinois Eastern Community Colleges consist of Wabash Valley College, Olney Central College, Lincoln Trail College, and Frontier Community College.

Appendix B. Top Thirty Most Frequently Occurring Tags from the Full I-Share Data, U of I Data, and U of I Records without Subject Headings

U of I Records without Subject Headings		Full U of I		Full I-Share	
Tag	Count	Tag	Count	Tag	Count
manga	77	photo	419	Bio	2,951
Action	73	history	338	Research	1,647
adventure	61	jkbnhs	327	psych	1,404
shounen	60	To Read	321	paper	1,279
supernatural	57	read	281	read	1,168
comedy	52	women	258	history	1,069
romance	50	paleo	253	book	1,065

U of I Records without Subject Headings		Full U of I		Full I-Share	
Tag	Count	Tag	Count	Tag	Count
Drama	39	Research	211	enviro	767
Historical	33	book	187	Philosophy	766
fantasy	32	music	182	film	594
demon	29	China	164	FYE	606
shoujo	29	wwd14	161	art	595
ghost	27	fiction	148	women	584
tournament	26	feminism	146	project	527
spirit	25	theory	131	To Read	512
fiction	22	manga	128	music	504
history	18	DigCand	124	Religion	480
To Read	18	ILRiver	124	theory	461
slice of life	16	Science	119	photo	460
book	15	Action	111	Education	443
canon	15	handbook	109	children's books portrayi	402
paranormal romance	15	social	109	english	394
ILRiver	14	design	108	social	380
Lesbian Pulp Fiction	14	Books	104	Oberg	366
literature	14	diss	103	design	359
read	14	Grinter	103	Thesis	339
Harem	13	comedy	102	jkbnhs	327
magic	13	shounen	102	fiction	323
HLM	12	Python	100	health	323
Literary fiction	12	Data	98	class	319