The Electronic Library

Comparing collaborative annotations on books between libraries and social community sites: A case study

Dan Wu Xiaomei Xu Wenting Yu

Received 24 September 2014
Revised 26 January 2015
Accepted 26 February 2015

Comparing collaborative annotations on books between libraries and social community sites

A case study

Dan Wu
School of Information Management, Wuhan University, Wuhan, China
Xiaomei Xu
Library, Zhejiang University of Water Resources and Electric Power, Hangzhou, China, and
Wenting Yu
School of Information Management, Wuhan University, Wuhan, China

Abstract

Purpose – Based on the study of overall situation of the tagging function in the provincial public libraries and library of major colleges and universities, this paper aims to examine the difference of tagging behaviour of its users in library and social community sites. The authors also want to understand the causes of a variety of annotation behavior in social community sites and libraries.

Design/methodology/approach – The authors collected all system log data of tags, comments and ratings users added in Wuhan University library, and then found the tags, comments and rating of corresponding books in Douban. Then, the authors did questionnaire survey to the Wuhan University students.

Findings – The authors found that the annotation service in the library is not perfect as that in social community site. Enthusiasm of users annotating books in the library is far less high than that on the social community sites. Lack of understanding of the annotation service is the main reason why users are not concerned or do not use the tagging service. But users have the needs of the organization of personal information in the library using tags.

Originality/value – This paper investigated the library users' behavior in the using library OPAC course and compared the difference of annotation behavior between library and social community site.

Keywords Academic libraries, Rating, Book tagging behavior, Comments, Social tags

Paper type Research paper

Introduction

Collaborative annotation, also known as *social annotation*, is the process by which web users use keywords or other natural language text to categorize, describe, comment or rate information resources online (Su *et al.*, 2010). Through sharing the annotated

This article is supported by National Program for Support of Top-notch Young Professionals in China.

The Electronic Library
Vol. 34 No. 2, 2016
pp. 178-195

0264-0473
DOI 10.1108/EL-09-2014-0171

information resources, Web users can collaboratively organize and retrieve the resources, obtain a more comprehensive and in-depth understanding of the resources and create special interest groups around the resources (Zauder *et al.*, 2007). Ever since the launch of the social cataloguing site LibraryThing in 2005, collaborative annotations on books have been increasingly popular on a variety of social community sites, such as Shelfari and Goodreads (Zhuang and Wu, 2012). On these sites, the related collaborative annotation activities include adding tags to books, commenting on books and providing ratings for books. For libraries, there are two benefits to expand and strengthen collaborative annotation services. First, collaborative annotation is a new method of information organization. Users can store bookmarks, add tags of their own choosing, rate and comment on books and designate individual bookmarks as public or private. Meanwhile, they can share the information by inviting others to view, comment, rate and offer feedback. The fact that the same resource has been tagged by more users allows for drawing connections between various users' collections and mutually tagged resources. The process thereby supports knowledge discovery, tag suggestion and insight into resource popularity, as well as interests and trends of users and communities (Zauder *et al.*, 2007). Second, collaborative annotation is user-centred. Thus, it can attract more users to participate actively and take advantages of library Web sites. With a growing number of annotations, information that users retrieve by tags and that libraries recommend to users can be more accurate. Libraries can use collaborative annotation tools, applications and folksonomies to enrich their annotation functions and expand traditional services.

Inspired by the success of collaborative annotations on social community sites, libraries started to provide such functions on their systems too. For example, among the surveyed 28 provincial public libraries and 112 academic libraries in China, it was found that about half (50 per cent) of the public libraries have the commenting function, 36 per cent have the rating function, but only 29 per cent have the tagging function. Some libraries provide external links to online book-related sites, such as Douban, Google, Baidu, Dangdang, Amazon, WorldCat, CNKI and so forth. Only seven libraries provided tagging, commenting, rating and external links.

As collaborative annotation is relatively new in library services, few studies have been conducted to examine collaborative annotations in libraries and library user annotation behaviours. Therefore, this article is motivated to conduct a case study on collaborative annotations in libraries and comparing them with those on social community sites. This study focused on using Wuhan University Library (WHUL) as one example from the library side and Douban as the other example from the social community site side. The research methods used involved log analyses and survey via questionnaires.

Literature review

Motivation for user tagging

Sen *et al.* (2006) proposed three factors that were likely to influence how people applied tags: people's personal tendency to apply tags based on their past tagging behaviours, community influence of the tagging behaviour of other members and the tag selection algorithm that chose which tags to display. Also, they summarized five kinds of motivations: self-expression, organizing, learning, finding and decision support. Through the investigation of different systems, Hammond *et al.* (2005) divided user

motivations according to who the tag was intended for and who the content creators into a "selfish" tagging discipline, where the users were primarily tagging their own content for their own retrieval purposes or a more "altruistic" tagging discipline, where the user was tagging others' content to allow others to retrieve. Golder and Huberman (2006) outlined seven individual types of tags based on tagging motivations, namely, identifying what or who, identifying what it is, identifying who owns it, refining categories, identifying qualities or characteristics, play and competition and task organizing.

Some research has been conducted based on tagging systems. Marlow *et al.* (2006) investigated the Flickr site and suggested that there were two kinds of tagging motivations: organizational and social. The authors discovered that a large part of the motivations and influences of tagging system users was determined by the system design and the method by which they were exposed to inherent tagging practices. Ames and Naaman (2007) performed in-depth, semi-structured interviews with 13 active Flickr users. A set of two-dimensional photo-annotation motivations was developed as a result: the first is "sociality" and the other "function". Organizational and communicational were two additional types of functions they identified. Some believed that the main motivation for Flickr users was social factors (Ames and Naaman, 2007; Marlow *et al.*, 2006), while others concluded that Del. ici. ous site users were motivated mainly by personal interests (Golder and Huberman, 2006; Hammond *et al.*, 2005). Bartley (2009) explored LibraryThing members' book-tagging purposes, the way tags were used and a comparison of member-assigned tags with terms in the corresponding MARC records. The results showed that LibraryThing members tag mostly for personal reasons, especially to manage their own collections, but they also tagged to assist others to find a book. Contrary to users of other popular tagging systems, such as Flickr, LibraryThing members did not perceive social networking as an important factor when assigning and using tags. The study also revealed that book tags by users varied greatly from MARC records. When tags overlapped with the terms used in MARC records, the majority of the matching occurred in the fields of 600s (subject access) and 245 (title statement). The analysis of other commercial Web sites (Amazon.com), including music Web sites (Last.fm), showed emerging social motivations for tagging which included opinion expression, performance and activism (Zollers, 2007). The various motivations for tagging in turn affected the nature of tags which themselves exhibit signs of sociality, as evidenced by their length and magnitude.

Studies of book-tagging behaviour of library users

Studies of tagging behaviour are mainly concentrated on a variety of tagging Web sites, such as Del. icio. us or Flickr. Tagging behaviour studies of library users are fewer and are mainly concentrated on the feasibility of the library's direct use of the tagging function and/or the opportunities of applying the social sharing sites within the library community. The existing research generally explored the feasibility of the application of tags to the library by the comparison of tags of social sharing sites and thesauri or subjects used in the library catalogue (Wu *et al.*, 2013). DeZelar-Tiedman (2011) obtained a sample of individual works of English and American literature from the twentieth and twenty-first centuries from a large academic library catalogue and searched for them on LibraryThing. She compared the subject headings and LibraryThing tags for match rates, their availability across both sets for various literary forms, and the terminology

used in the tags versus the controlled vocabulary of the bibliographic records. In addition, the usefulness of available LibraryThing tags for the library catalogue records that lacked subject headings was evaluated. Lu *et al.* (2010) compared the LibraryThing tags for a group of books and the library-supplied subject headings for the same books. The difference between users and cataloguers in resource description showed that user tags can enhance subject access to library materials, but cannot entirely replace controlled vocabularies. To explore the feasibility and obstacles of applying tags to library records, Rolla (2009) compared the LibraryThing tags for a group of books and the library-supplied subject headings for the same books and found that the tags could improve the availability of library collections, except for ones irrelevant to the subject headings that might hinder their application to the library bibliographic systems. Thomas *et al.* (2009) found that social tagging did indeed augment Library of Congress Subject Headings (LCSH) by providing additional access to resources and a hybrid catalogue combining both LCSH and a folksonomy results in richer metadata and be stronger than the sum of its parts, giving patrons the best of both worlds in terms of access to materials. Other methods, such as NISO standards, were used to analyse and evaluate tags (Spieteri, 2007). The possibilities of applying social sharing sites to the library are also discussed (Mendes *et al.*, 2009; Rethlefsen, 2007).

While the above studies may show the application of social features in a library setting can contribute a lot, it is also important to examine the reality of these benefits by conducting comprehensive usability studies. Kakali and Papatheodorou (2011), for example, evaluated OPACIAL by a technology acceptance experiment, in which 20 users (postgraduate students and faculty members) used all its functionalities for a week, inserted more than 500 tags and, finally, were interviewed to assess the system usability and usefulness. One of the important findings of the interviews was that the users, in general, consider tagging functionalities useful, though they were sceptical when browsing the tag cloud and were also fearful of its constantly expanding size. Conversely, Spieteri and Tarulli (2012) presented a comparison of two social discovery systems used in two Canadian public libraries by the analysis of four months' worth of log analysis, and the results showed that the social features were underused. Both systems provided more options by which user-generated content could be accessed and searched. They also emphasized the importance of the extent to which the patrons were motivated to add tags, reviews or ratings.

Overall situation of the tagging function in the provincial public libraries and library of major colleges and universities in China

In practice, the number of libraries that use annotation functions – in particular the tagging function – is still relatively small. The forms and quantities of application of the annotation function vary a lot; there is no uniform standard for using social annotation functions. For example, some libraries only provide one annotation function from among tagging, commenting or rating, whereas other libraries use all three functions. Some libraries only allow users to add tags, but do not permit searching, browsing or organizing tags. In addition, most of the libraries that provide the tagging function (i.e. libraries using the Ex Libris or InterLib systems) allow organization and management of tags and allow users to organize their own tags into a "personal library", but only a few libraries allow organization and management of personal comments.

Feng and Wu (2011), based on the relationship within social book annotation Web sites, classified the application modes of social tagging in libraries into four styles: combination style, embedded style, hybrid style and extended style. This study identified that the most widely used mode in China is the combination style, providing the library catalogue with both the tagging functions and links to external resources. The second most widely used is the embedded style, with a total of 33 academic libraries and 13 public libraries using this mode. Two libraries, Jinan University Library and the University of Electronic Science and Technology Library, use the hybrid mode. In addition, Jinan University Library provides its users with Google browser plug-ins so that users can easily search the collection of books in the library OPAC using the Google Toolbar. The plug-ins provided by the University of Electronic Science and Technology library can embed the bibliographic function of the library into the Douban network, allowing users to directly search books in the school library's collection in the Douban book (2012) search page. In addition, as far as the aspect of using the annotation to enhance the interactions with readers, most of libraries use popular tags, hot comments/hot reviews/top ten reviews and top rank books. But few libraries provide the comment reply function. Our study found that only the Anhui University of Technology Library provides this feature. Therefore, the interaction using annotations between libraries and users is weak. As for the book recommendation function, few libraries provide personalized recommendations based on the users' own tags, comments and collection, but base these recommendations on some other popularity measures, such as top hits, popular evaluations/reviews, popular books and so forth.

Research design

Research questions

First, it is important to make it clear whether and how such services in libraries differ from those on social community sites. Therefore, the first research question is:

RQ1. What are the characteristics of library users' collaborative annotation behaviours? What are the differences between these behaviours and that of the users in the online social community sites?

Previous work shows that collaborative annotations are underused in libraries (Han and Liu, 2010). This study, therefore, focused on the barriers that prevent library users from actively using collaborative annotations. Specifically:

RQ2. What are the barriers that prevent users in libraries from using collaborative annotations? Are these barriers different to those on online social community sites? What can libraries learn from social community sites on overcoming these barriers?

Research methods and data collection

The study in this paper involved mixed research methods that are appropriate for answering each research question. For the first research question, to obtain in-depth study results, a case study approach was adopted as the research method, and the WHUL (http://en.lib.whu.edu.cn/web/en/home) was selected as the sample case on the library side. As an academic library of a top research university in China, WHUL has been providing collaborative annotation functions, such as tagging, commenting and rating, in their library OPAC system since 2009. The other sample case selected on the

social community site side was Douban (www.douban.com) which is the largest online social network and annotation site in China. Its book-related section (simply called "Douban book"; http://book.douban.com) has a long-serving history with over several million users.

Because of the authors' unique relationship with WHUL, the library's system logs were obtained for this study. These logs cover annotation activities from the library's launch day in 2009 up to 31 March 2012. All the information was extracted about the books annotated by library users during the log period. The WHUL system logs were run through Excel to extract the descriptive statistics. Using the ISBN number, title, publishers and other related information for these annotated books, the corresponding annotations were collected (e.g. tags, comments and ratings) on Douban. This task was conducted between 6 and 11 April in 2012. Table I summarizes the annotation information from the WHUL logs and that of their corresponding books on Douban. Apparently, almost all of the annotated books in the library are popular books that have corresponding annotations on Douban as well.

For the second research question, a questionnaire survey method was used. Among all the valid questionnaires, WHUL users were specifically identified who pointed out that he/she had annotation experience on the library system or social community sites on the survey, to further explore the motivations and obstacles in users' annotations. The focus of the questionnaire was about the users' collaborative annotation behaviours and motivations on the WHUL and social community Web sites. The design of the questionnaire was influenced by the work of Bartley (2009), as well as an early pilot survey. The questionnaire consists of three parts, namely, background checks, the online book annotation behaviour survey and the library book annotation behaviour survey. The questions in the questionnaire consist of single choice, multiple choices and Likert-scale questions with 1-5 indicating low to high (1 = low and 5 = high). On 1 June 2012, 400 copies of the questionnaire were distributed at the WHUL. The results were analysed using SPSS and Excel.

Participants

The 400 survey questionnaires distributed in WHUL were for the second research question. Before handing out the questionnaire, the WHUL users were asked whether he/she knew about collaborative annotation and only those who were familiar with it were chosen as participants. In total, 357 participants from multiple schools in Wuhan University returned completed survey questionnaires. The distributions of the participants' majors were: science 23.2 per cent, engineering 15.1 per cent, social science 41.2 per cent, humanities 20.2 per cent and medical science 0.6 per cent. There were 279 undergraduates (78.2 per cent), 74 master's students (20.7 per cent) and 4 doctoral

Table I. Annotations obtained from the Wuhan university library and Douban

Number of annotations	Tags	Comments	Ratings	All three annotation forms
Number of books containing the specified annotation form in the WHUL logs/number of corresponding books found on Douban	270/269	124/118	278/254	10/10
Total number of the specific annotations found in the WHUL logs	315	139	337	

students (1.1 per cent). The majority (71.4 per cent) of the participants were on the Internet 2 to 4 h per day, 16 per cent were online for 1 h or less, 10.6 per cent were between 5 and 7 h and 2 per cent were more than 8 h. Of the participants, 44.5 per cent were male and 55.5 per cent were female. The degree of understanding tagging, commenting and rating were 3.1, 3.3 and 3.1, respectively. This means that their familiarity with collaborative annotation was passable, and there was no great difference in their understanding of the three different annotation actions.

Results

Analysis of Wuhan University Library logs

Two methods were used to classify the annotated books in the WHUL logs. The first classification adopted the Chinese Library Classification (CLC) to group the books into 22 categories from A to Z. This classification helped to examine the books from the view point of a traditional library classification scheme. As shown in Figure 1, CLC class T (Industrial Technology) contains the highest number of books, followed by class I (Literature), class R (Medicine and Health Sciences), class O (Mathematics, Physics and Chemistry) and class F (Economics). When the books in class T were further divided into subcategories, it was found that subclass TP (Automation, Computer Engineering) had the most annotated books (64), followed by TN (Electronic Engineering, Telecommunication Engineering) with 31 books. These results show that the majority of the annotated books are focused around several major categories, which are related to the academic studies in the university. However, books on literature, which might be related to the non-academic side of university life, are also popular targets for annotation.

The second classification was to examine the annotated books from the view point of the online social community site. However, Douban only provides a simple two-layer classification based on tags. Therefore, the classification schema used on Amazon.com and Dangdang.com (the largest Chinese online bookseller) were examined, and the following online classification scheme was proposed: Textbooks (teaching materials), teaching supplements (teaching aid books/reference books), professional and technical, best sellers (literature/art), popular science (business/history/politics, etc.) and lifestyle (entertainment/sports/travel/health/children, etc.). Figure 2 shows the distributions of the annotated books based on this classification schema. Almost half of the books are in the professional and technical class. The next two are in the best sellers and textbooks classes, each of which accounts for 16 per cent, respectively. The smallest class is lifestyle. This shows that most books annotated by the university library users are related to their academic activities.

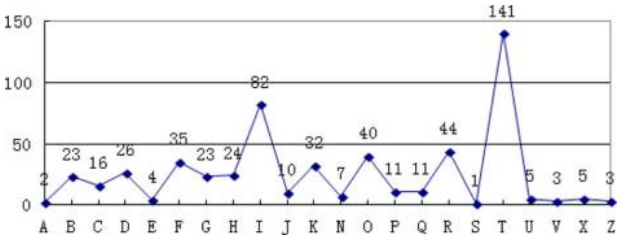

Figure 1.
Number of annotated WHUL books classified according to CLC

Comparison of tagging behaviours between Wuhan University Library and Douban

Based on the annotated books found in both WHUL and Douban, a comparative study between the tagging behaviours on these two sites was performed. The comparison focused on several aspects of tagging behaviours with the hope of providing comprehensive reviews.

Numbers of tags. It is straightforward to obtain the number of tags for each book in WHUL, but it is rather more complicated for Douban. Douban uses a number in parentheses after a tag – for example, psychology (245) – to indicate the usage frequency of the tag. However, it only displays the top eight most frequently used tags (hereafter referred to as commonly used Douban tags in this paper) with an indication of the total number of tags for a book. Therefore, as shown in Table II, the top row is for tags obtained from WHUL and the next two rows are for tags obtained from Douban, which shows both commonly used and all tags from Douban. The data show that the number of tags obtained from Douban is much higher than that from WHUL, both in terms of average tags per book and total number of tags. In addition, probably due to the low number of annotations, the maximum and minimum numbers of WHUL tags per book is five versus one, whereas the same set of books in Douban can have as many as 1,635 tags and yet as few as zero tags. The standard deviation of commonly used Douban tags is 3.62, which is higher than that of WHUL (0.55) in an order of magnitude. Of course, due to its popularity, the average number of all Douban tags can reach to 44.04 tags per book and change a lot from one book to another (with 93.14 as its standard deviation). Tagging on community-based social sites is much more popular than on a library site.

Types of tags. Following the work by Cantador *et al.* (2011), tags were divided into four categories: content-based, context-based, subjective and organizational. *Content-based* tags describe the content of the book, such as "food" and "computer". *Context-based* tags provide contextual information about the books, such as the tag "Japan" used to describe the place a book is published and " 2005" denoting the publication year of a book. *Subjective* tags present the personal view (e.g. like, cute, etc.) and the quality of books. An example of a subjective tag for the book *C++ Primer Plus* (5th edition, Chinese version) is "classic". *Organizational* tags include personal uses and tasks, as well as the tags that have the personal reference value, such as "to read" and "my favourite collection". As shown in Table IV, the majority of tags in both the WHUL

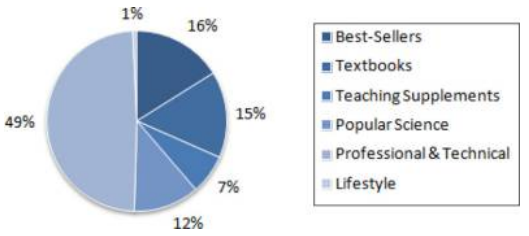

Figure 2.
Distributions of annotated WHUL books according to the proposed online classification scheme

Tags	Total	Average	Maximum	Minimum	SD
WHUL tags	314	1.17	5	1	0.55
Commonly used Douban tags	964	3.57	8	0	3.62
All Douban tags	11,760	44.04	1,635	0	193.14

Table II.
Number of tags on WHUL and Douban (with duplicate tags removed)

and Douban collections are content-based tags. Thus, annotating the book content is the main purpose for users to add tags. Interestingly, although context-based tags are the smallest on WHUL, they are the second largest category on Douban. Further analysis shows that most context tags are related to countries or based on the period and/or location(s) described in the book. This information is often available in a library's bibliographic records, but may not be available in Douban's records. This might be the reason why these tags are popular on Douban but not on WHUL. Organizational tags are ranked second on WHUL, but are last on Douban. It seems that the demands for personal information organization often trigger users to add tags to library books, whereas such demands are relatively weak on Douban. This is probably because Douban already organizes books into "want to read", "reading" and "read before" before tags can be added to books. Such a service is not available on WHUL. Subjective tags are the third most popular category on both WHUL and Douban. Their numbers and proportions are relatively low because subjective tags are added based on the users' level of feeling towards the books (e.g. love or hate), making them more or less like comments (Table III).

The overlaps were calculated for the tags from WHUL and Douban and mapped into the types of tags mentioned in the above classification. As shown in Table IV, almost all the overlapping tags are content-based tags (48 of 50). This is not surprising, as content is probably the most important aspect when users on WHUL and Douban annotate the books. Because WHUL has a smaller number of tags, the percentage of the overlapping tags on WHUL is much higher than that of Douban. Because of the relatively low percentage on both WHUL and Douban tags, library users and Douban users do not often use identical tags for the same books even when they try to describe the content of the same book. There is almost no overlap on the other three types of tags at all.

Forms of tags. Following the classification of tags based on their forms in the literature (Yi and Chan, 2009), tags were classified into tags in Chinese words (such as

Table III. Types of tags on WHUL and Douban

Tags	Content-based (%)	Context-based (%)	Subjective (%)	Organizational (%)
Number of WHUL tags (%)	210 (66.88)	5 (1.59)	27 (8.60)	72 (22.93)
Number of Douban commonly used tags (%)	843 (87.45)	63 (6.53)	31 (3.22)	27 (2.80)

Table IV. Overlaps of tags from WHUL and Douban

Types of tags	No. of overlap tags	Percentage of WHUL tags	Commonly used Douban tags (%)
Content-based tags	48	15.2	5.0
Context-based tags	2	0.6	0.2
Subjective tags	0	0	0
Organizational tags	0	0	0
Total	50	15.8	5.2

"心理学"; translated as "psychology"), tags in Chinese phrases (such as "优先看"; translated as "first to read"), tags in Chinese sentences (such as "机器智能是未来生物之后一大热门"; translated as "machine intelligence is the future, a hot topic after biology"), tags with non-Chinese words (such as "biochemistry"), tags with symbols only (such as "TP31/Sh10") and hybrid tags (such as "C语言"; translated as "C language"). Table V shows the distributions of these tags. It can be seen that the proportions of tags in Chinese words and tags in Chinese phrases are the highest on both WHUL and Douban, whereas the proportion of tags in Chinese sentences are the lowest on both collections. However, the number and the proportion of tags with symbols only on WHUL are both significantly higher than that on Douban. Because tags with symbols are often only meaningful to the individual who creates them, their usage is often low. Therefore, this can be problematic in Douban, as Douban only displays the top eight most frequently used tags. However, WHUL shows all the tags. Maybe this is the reason that tags with symbols only are much more common on WHUL than on Douban. There is no significant difference between the two collections on other forms of tags such as hybrid tags and tags with non-Chinese words.

Comparison of comments and ratings between Wuhan University Library and Douban Comments. Table VI shows the descriptive statistics of the comments collected from WHUL and Douban. There are great differences between the two sets of comments. No matter whether it is the total number, the average, the maximum or the standard deviation, WHUL comments are far less than those on Douban, which indicates the much greater number of comment activities on Douban. The much bigger standard deviation on Douban comments also indicates that the number of comments on each book varies much more widely on Douban than on WHUL. When examining the comments in details, the Douban comments are often long, with length ranging from a

Forms of tags	WHUL tags: no./percentage	Commonly used Douban tags: no./percentage
Tags with Chinese words	84/26.8	566/57.7
Tags with Chinese phrase	148/47.1	299/31.0
Tags with Chinese sentence	3/1.0	9/0.9
Tags with non-Chinese words	35/11.1	68/7.1
Tags with symbols only	35/11.1	12/1.2
Hybrid tags	9/2.9	10/1.1

Table V. Forms of the tags from WHUL and Douban

Comments	Total	Average	Maximum	Minimum	SD
Library comments	133	1.1	3	1	0.46
Douban comments	9,639	77.7	1,998	0	257.42

Table VI. Comments of books from WHUL and Douban

few dozen to several hundred words, and even up to several thousand words. Some comments on Douban are actually well-written book reviews, which provide deep understanding of the content and further related readings. In contrast, WHUL comments are often short, and the majority only contain a few dozen words. However, these WHUL comments have very diverse content types, that is, book content, quality evaluation, book review and book recommendation. One interesting issue found within the WHUL comments is that some WHUL comments actually contain problems encountered in the use of the library ("CD with the book"), the recommendations of the library ("I hope the school ordered several more") and even socialization with other readers. This demonstrates that users do have the need for communicating with the library and other users on specific issues around the books, and comments are viewed as the focused and in-context communication channel. This raises a question for the library to fully take advantage of the comment services for better understanding of users' needs, for better interaction with their readers and for better resolution of users' encountered problems. Thus, the comment replying service in Douban may shed light on a solution to this question. That is, libraries can provide services, such as comment creation, reply, share, recommendation, subscription, interest group building and so forth.

Ratings. Both WHUL and Douban have many ratings on the books. The number of ratings made on Douban is 492,020 to 492,570 (the reason this is a range is because Douban would not give an exact number when there are less than 10 ratings to a book; it just simply states "less than 10"), far greater than that on WHUL, whose number is 337. As the rating consistency between those on WHUL and those on Douban is of interest, a Kappa test was performed on the two sets of ratings. The rating scores were first normalized on Douban from a ten-point to a five-point scale as on WHUL, and then the SPSS software was used to generate a Kappa coefficient value. The score is merely 0.031, indicating that the consistency is low. However, as $p > 0.05$, this result does not have adequate confidence.

Comparison of library users' attention on book annotations

As part of the study, 357 users of the WHUL were surveyed. Table VII shows the users' attention to the annotations in the social community site and in the library's OPAC. It can be seen that, on the social community sites, users' concerns are mainly with bibliographic information, comments, ratings and book recommendations, whereas they have less interest in tags and "groups, friends and forums". This overall seems to indicate that the users are more focused in their own reading interests, which suggests that social needs are relatively low on these sites. When looking at the notes on the

Table VII. Concern for various items on social community and library sites

Social community sites		WHUL		*t*-test
Information items	Means	Information items	Means	*P*-value
Tags	2.77	Tags	1.78	0.000
Comments	3.71	Comments	2.04	0.000
Ratings	3.48	Ratings	1.97	0.000
Bibliographic information	3.80	Bibliographic information	4.06	0.005
Book recommendations	3.21	Book location	4.29	–
Group, forum and friends	2.57	Network summary	2.99	–
		Douban comments	2.12	–

library OPAC records, the users paid high attention to the general bibliographic information and the location of the book in the collection location. There is some attention to the abstract of the books. Beyond that, the rest of the items, such as tags, ratings and comments, are all given pretty low attention.

Comparing the data collected from the library and from the social sites, the attention on comments and ratings on the social community sites is also higher than that on the WHUL site. People's attention to the bibliographic information is lower on the library's site. Statistical tests analysing the differences between users' attentions on various records on social sites and that on the library site show that there are significant differences on the comments, rating evaluations, bibliographic information and tags ($p < 0.05$). This shows that users on social community sites are more concerned about tags, comments and ratings, whereas users pay more attention to bibliographic information in the library's OPAC.

Motivations and obstacles in users' annotations

As part of the survey of those 357 participants, answers were collected as to their motivations for annotations on social community sites. Because 61 of the 357 respondents had not added any tags, comments or ratings on social community sites before, they were removed from further study. The motivation analysis below was based on the remaining 296 respondents.

Motivations for annotations on social community sites. As shown in Figure 3, "recommending books and evaluating quality" is the most popular motivation, accounting for 56.95 per cent. The next three, which have roughly the same popularity, are "explaining content of the resources", "recommending books to themselves" and "organizing personal book collection". The motivations for helping others, such as "helping friends, groups or the site" and "helping others to search for books" are the least popular ones. This shows that users on social community sites are not really for social purposes at all. They just want to work on their own tasks.

Obstacles for annotations on social community sites. Further analysis of the obstacles that prevent users from adding annotations on the social community sites was conducted. As shown in Figure 4, the top four obstacles were:

(1) "not interested in annotations";

(2) "not understanding annotations";

Figure 3.
Motivations for users' annotations on social community sites

(3) "annotations are too complex and too difficult to use"; and

(4) "self-annotations are not accurate".

The least mentioned obstacles were "not read often" and "not helpful for their own lives and learning". This result shows that most users view annotations as somewhat helpful for their own studies and tasks. A few users also stated their obstacles as "too much trouble" or "need to register to add annotations", which were treated as evidence that annotations are difficult to use. All these show that users still need lots of help in understanding the importance of annotation and in using annotations.

Obstacles for not viewing annotation results on the Wuhan University Library site. As part of the study, WHUL's system logs showed that the total number of students who added tags, comments or ratings to books in the library is 536. This number is smaller than that of online sites, so the obstacles for annotations in the library site were studied as well. Figure 5 shows that users still do not really know what the annotation function is. The two most common obstacles are related to the misunderstanding of annotation:

(1) "only to retrieve the physical location of the book"; and

(2) "did not know that the library has such services".

Some less common obstacles include "did not find the wanted book on library site" and "had no interest or time to be concerned with this service". All these demonstrate again that the library should promote their annotation service more and provide better training to the users.

Obstacles for not adding annotations on the Wuhan University Library site. Figure 6 shows the obstacles that prevent users from adding annotations on the WHUL site. The top two reasons are again "did not know the library has such services" and "library site is only used for retrieving the physical location of the book". It seems that the issues are the same – users do not know of or misunderstand the annotation function.

Overall, the above results demonstrate that no matter whether it is on social community sites, such as Douban, or library sites, such as WHUL, the main motivations are related to users' own tasks, and the obstacles for using annotations are that the users do not really know about or misunderstand annotations. It is important to promote book annotations before its wider usage is achieved.

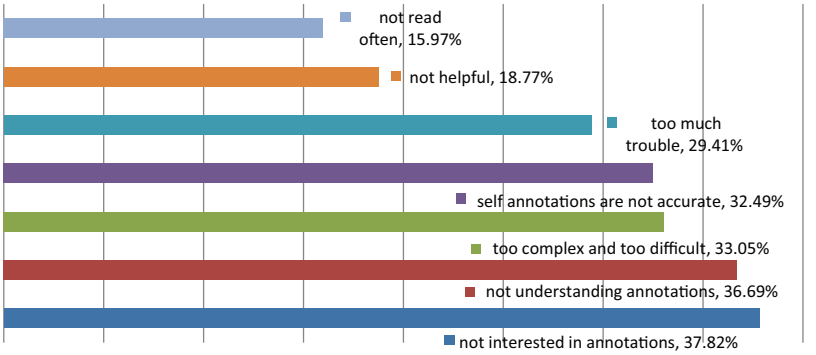

Figure 4.
Obstacles that prevent users from adding annotations on social community sites

Users' expectations of the annotation service in libraries

In the survey of the 357 respondents, their expectations of the annotation service in libraries were asked. Participants were asked to rate a list of expectations with scores ranging from 1 to 5 (1 being the lowest and 5 being the highest). Table VIII shows the average scores of the expectations. All annotation services received a certain degree of demand from the participants, and the differences among their expectations were not large. Recommending books, retrieving tags and organizing tags are the three most demanded services, which again show that users are more interested in using book annotation in libraries to help them find books. The social aspects of book annotation, such as socializing, commenting and rating, are slightly less expected. They have the least demands on seminars or other open events about annotation services.

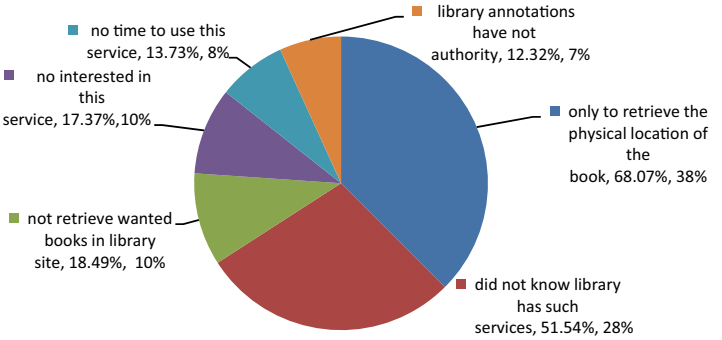

Figure 5.
Obstacles that keep users from viewing annotations on the WHUL site

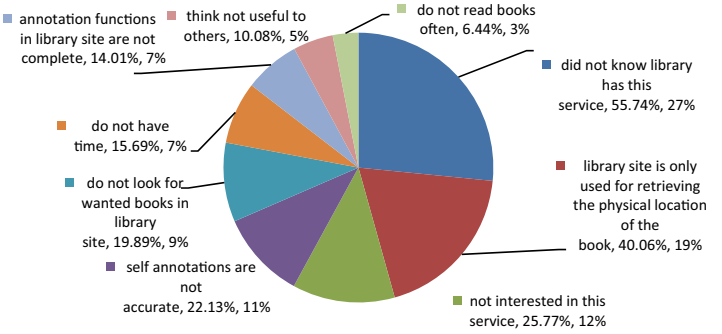

Figure 6.
Obstacles preventing users from adding annotations on the WHUL site

Statistical items	Organizing tags	Retrieving tags	Recommending books	Socializing in friends, groups, and forums	Commenting and replying comments	Rating	Promoting and holding seminars
Mean	3.4622	3.7983	3.8235	3.4006	3.3333	3.4146	3.2241
Standard deviation	1.23245	1.09066	1.12914	1.12893	1.17240	1.23462	1.22941

Table VIII.
Users' expectations of the annotation service in libraries

Discussion

Summarizing based on the above result analyses, five observations on the differences between users' annotation behaviours on library and social community sites can be seen. First, the numbers of book annotation activities, such as providing tags, comments or ratings, in library sites were relatively small compared to the numbers from the social community sites, and the numbers of tags and comments on each book in library sites were small too. This is true for all the library users who performed annotations, which is reflected by the small standard deviation. It seems that library sites have not been identified as a place for annotating books, even though there is a much longer history for libraries to have book-related information than social community sites.

Second, content-related tags and subjective tags were very popular on both library and social community sites. However, organizational tags were more common on the library sites and rare on the social community sites. Contextual tags, on the other hand, had the opposite distribution. These seem to indicate that library site users have more intentions to organize books, which is in line with most library missions.

Third, although the proportion and the ranking of various types of tags were the same between the library and the social community sites, symbol tags had higher numbers and percentages on the library sites than that of social community sites. This shows that the authority control of the tags in the library was not as good as that on the social community sites, probably because less people used the service on the library site. Additionally, fewer library users wrote comments, and even when they wrote comments, the contents of the comments were often simple. Some of the comments were about library services or suggestions to the library. In contrast, the comments on the social community site were longer, the content was richer, the coverage was wider and the quality was generally higher. It shows that library sites need to make more efforts to promote book reading guidance and book recommendation functions on the library sites.

Finally, the main obstacles for not using book annotation functions on the social community site were related to the users not being interested or not knowing about annotation, whereas the main reasons for users not annotating or not seeing the annotation on the library sites are that they do not know about the annotation service or they do not know of the OPAC function. Therefore, promoting annotation services would help.

The results show that, on the one hand, users do not have much expectation of the annotation service in the library, but, on the other hand, the users' lack of understanding of the annotation service is the main reason that users do not use the annotation service. At the same time, the results show that users have demands of various kinds of annotation services. Therefore, it is critical to promote annotation services in the library, and that, particularly, the influences can be shown as having two aspects:

(1) *Clarifying the concept of user annotations*: One reason that users do not annotate books is because they worry about the accuracy of their own annotations. As a user-centric Web 2.0 service, what users should pay attention to is how their annotations reflect their own understanding, and how the annotations help them and others in organizing and retrieving books. Users should be taught to have fewer concerns about accuracy.

(2) *Introducing and training on libraries' annotation services*: A constant message obtained from the study is that library users do not understand the annotation functions provided by the library. Thus, introducing the annotation services and providing training on them are essential. Therefore, libraries should make use of

a variety of ways and forms to introduce annotation services in a simple, clear and straightforward way, thereby enhancing the users' understanding of the service and hopefully increasing their frequency of annotation.

Another one of the main reasons why users do not add or view the annotation that found in the survey is that the user is not interested in this service. The author believes that annotation services are on the rise as a new service and have not yet gained great popularity. Part of the reason why users are not interested in this service is that users do not yet understand the benefits of this service for users. Therefore, libraries should use various channels and methods to introduce annotation services in simple and concise methods, so as to improve users' adoption of annotations and their usage frequency of the functions.

Conclusion

Annotation functions have many advantages, but perhaps users will recognize this since social community sites have done so well in this regard because users can realize their own needs on these sites, so it may not require libraries to repeat this function. However, user behaviour analysis revealed that most annotated books in the library are those users often use or are interested in, and also that frequently annotated books are related to the professional or to learning. This shows that library user's annotated object is unique and is consistent with the characteristics of university library users. Therefore, social community sites cannot completely replace the library. However, users' annotated positivity in the library is significantly lower than that on social community sites; one of the reasons is that the library does not know enough about the user's needs. Whether the user is willing to use the service of the library, is familiar with the service, and can use this service decides whether it is necessary and ongoing effects of good or bad. Therefore, libraries should carry out detailed investigations on their users to better understand whether users are willing to use this library service and, perhaps through a variety of seminars or microblogging platforms, libraries can propagandize the book annotation function and conduct user education. As found in the survey, Wuhan University students have less concern about friends, groups or forum in the annotation course, and the expectations of this function is lower than that of other functions. The library can be first to carry out other annotation services, and later when conditions permit and the demand of readers increase, the library can then carry out the services.

At the same time, libraries should improve annotation capabilities and annotation functions provided by the library should be easy to use. For example, the library should organize tags users add to their personal library, to achieve the classification and retrieval of a personal collection of books, as well as tag retrieval and personalized book recommendations. The more perfect the annotation function becomes, the higher the probability of users taking advantage of the annotation function will be. The frequency of using the annotation service and the annotation effect influence each other. For example, the greater the number of users using the annotation service is, the more accurate the annotation result will be, and the books recommended for the user will be more accurate. The higher the efficiency of the annotation service is, the more patrons are likely to use it and become consistent users of the service.

In short, libraries can learn from the successful experience of book-related social community sites, with its own characteristics, and user-centric and actively provide users with a richer, more targeted book annotation service.

References

Ames, M. and Naaman, M. (2007), "Why we tag: motivations for annotation in mobile and online media", *Proceedings of the SIGCHI Conference on Human factors in Computing Systems*, *ACM Press, New York, NY*, pp. 971-980.

Bartley, P. (2009), "Book tagging on LibraryThing: how, why, and what are in the tags?", *Proceedings of the American Society for Information Science and Technology*, Vol. 46 No. 1, pp. 1-22.

Cantador, I., Konstas, I. and Jose, J.M. (2011), "Categorising social tags to improve folksonomy-based recommendations", *Web Semantics: Science, Services and Agents on the World Wide Web*, Vol. 9 No. 1, pp. 1-15.

DeZelar-Tiedman, C. (2011), "Exploring user-contributed metadata's potential to enhance access to literary works", *Library Resources & Technical Services*, Vol. 5 No. 4, pp. 221-233.

Douban Book (2012), Available at: http://book.douban.com (accessed 8 June 2012).

Feng, Q.R. and Wu, D. (2011), "Application patterns of social tagging in libraries", *Library and Information Service*, Vol. 56 No. 21, pp. 88-92.

Golder, S.A. and Huberman, A. (2006), "Usage patterns of collaborative tagging systems", *Journal of Information Science*, Vol. 32 No. 2, pp. 198-208.

Hammond, T., Hannay, T., Lund, B. and Scott, J. (2005), "Social bookmarking tools (I)", *D-Lib Magazine*, Vol. 11 No. 4, available at: www.dlib.org/dlib/april05/hammond/04hammond.html (accessed 6 February 2012).

Han, Z.P. and Liu, Y.Q. (2010), "Web 2.0 applications in top university libraries", *Library Hi Tech*, Vol. 28 No. 1, pp. 41-62.

Kakali, C. and Papatheodorou, C. (2011), "The exploitation of social tagging in libraries", *First Workshop on Digital Information Management*, *Corfu*, 30-31 March, available at: http://eprints.rclis.org/15850/1/07.Kakali.pdf (accessed 8 June 2012).

LibraryThing (2012), available at: www.librarything.com/zeitgeist (accessed 8 June 2012).

Lu, C.M., Park, J. and Hu, X.H. (2010), "User tags versus expert-assigned subject terms: a comparison of LibraryThing tags and library of congress subject headings", *Journal of Information Science*, Vol. 36 No. 6, pp. 763-779.

Marlow, C., Naaman, M., Boyd, D. and Davis, M. (2006), "HT06, tagging paper, taxonomy, flickr, academic article, to read", *Proceedings of the 17th Conference on Hypertext and Hypermedia*, *ACM Press, New York, NY*, pp. 31-40.

Mendes, L.H., Quiñonez-Skinner, J. and Skaggs, D. (2009), "Subjecting the catalog to tagging", *Library Hi Tech*, Vol. 27 No. 1, pp. 30-41.

Rethlefsen, M.L. (2007), "Tags help make libraries Del.icio.us: social bookmarking and tagging boost participation", *Library Journal*, Vol. 132 No. 15, pp. 26-28.

Rolla, P.J. (2009), "User tags versus subject headings: can user-supplied data improve subject access to library collections?", *Library Resources & Technical Services*, Vol. 53 No. 3, pp. 174-184.

Sen, S., Lam, S.K., Rashid, A.M., Cosley, D., Frankowski, D., Osterhouse, J. and Harper, F.M. (2006), "Tagging, communities, vocabulary, evolution", *Proceedings of the 2006 20th Anniversary Conference on Computer Supported Cooperative Work (CSCW '06)*, *ACM*

Press, New York, NY, available at: http://dl.acm.org/citation.cfm?id=1180904 (accessed 8 June 2012), pp. 181-190.

Spieteri, L.F. (2007), "The structure and form of folksonomy tags: the road to the public library catalog", *Information Technology and Libraries*, Vol. 26 No. 3, pp. 13-25.

Spieteri, L.F. and Tarulli, L. (2012), "Social discovery systems in public libraries: if we build them, will they come?", *Library Trends*, Vol. 61 No. 1, pp. 132-147.

Su, A.Y.S., Yang, S.J.H., Hwang, W.Y. and Zhang, J. (2010), "A web 2.0-based collaborative annotation system for enhancing knowledge sharing in collaborative learning environments", *Computers & Education*, Vol. 55 No. 2, pp. 752-766.

Thomas, M., Caudle, D.M. and Schmitz, C.M. (2009), "To tag or not to tag?", *Library Hi Tech*, Vol. 27 No. 3, pp. 411-434.

Wu, D., He, D.Q., Qiu, J., Lin, R.N. and Liu, Y. (2013), "Comparing social tags with subject headings on annotating books: a study comparing the information science domain in English and Chinese", *Journal of Information Science*, Vol. 39 No. 2, pp. 169-187.

Yi, K. and Chan, L.M. (2009), "Linking folksonomy to library of congress subject headings: an exploratory study", *Journal of Documentation*, Vol. 65 No. 6, pp. 872-900.

Zauder, K., Lazi, J.L. and Zorica, M.B. (2007), "Collaborative tagging supported knowledge discover", *Proceedings of the 29th International Conference on Information Technology Interfaces*, *IEEE Press, New York, NY,* pp. 437-442.

Zhuang, X.Z. and Wu, D. (2012), "Comparison and analysis of book tagging systems", *Library and Information Service*, Vol. 56 No. 5, pp. 73-77.

Zollers, A. (2007), "Emerging motivations for tagging: expression, performance, and activism", paper presented at WWW 2007, Banff, 8-12 May, available at: www2007.org/workshops/paper_55.pdf (accessed 8 June 2012).

About the authors

Dan Wu is Professor of the School of Information Management, Wuhan University. Her research interests include information retrieval, user behaviour and information organization. Dan Wu is the corresponding author and can be contacted at: woodan@whu.edu.cn

Xiaomei Xu is a Librarian of the Library of Zhejiang University of Water Resources and Electric Power. She graduated from the School of Information Management, Wuhan University. Her research interests focus on user behaviour.

Wenting Yu is a master's student of the School of Information Management, Wuhan University. Her research interests focus on user behaviour.