A Quantitative Analysis of Collaborative Tags:

Evaluation for Information Retrieval—a Preliminary Study

J. Gelernter

School of Information and Library Studies
Rutgers University
4 Huntington Street New Brunswick, N.J. U.S.A.
gelern@rci.rutgers.edu

Abstract—**Collaborative (or social tagging) options are being added to many database catalogs on the assumption that not only those who assign tags but also those who use the catalog find such tags beneficial. But no quantitative analyses of collaborative tags exist to support this assumption. Based on questionnaires mixing collaborative tag clouds from http://www.LibraryThing.com and controlled Library of Congress Subject Heading (LCSH) strings from the Library of Congress catalog http://catalog.loc.gov, it was found that controlled vocabulary terms are selected above collaborative terms; that the string format is preferred to the cloud; that strings appear to "perform" better in terms of reflecting book content; and that it is important to users that recall is high (where uncontrolled vocabulary retrieval is generally low). Results were found to be dependent upon particulars of tag cloud or string. The outcome indicates that catalog users would derive fewer information retrieval benefits from the current form of collaborative clouds than from the staid strings of the Library of Congress Subject Headings**.

Keywords-collaborative tags; social tagging; tag cloud; recall; information retrieval

I. INTRODUCTION

A panel comprised of creators and commentators on social tagging systems met during the Human Computer Interactions conference during April 2007 in Montréal, Canada. They considered questions such as how the design of tagging systems affects user participation, and what are expected patterns of change over time as systems scale [7]. Information retrieval benefits of tagging were assumed. But the debate about the relative merits of controlled versus uncontrolled vocabularies for information retrieval continues [16], [22].

Uncontrolled terms in many earlier experiments were in the context of computer-generated indexing; here the uncontrolled terms are generated not by computer but by social taggers. Enthusiast-generated tag clouds already have been absorbed into a few academic catalogs such as that of the University of Pennsylvania. A beta version appears at OCLC http://orlabs.oclc.org/Identities/. This study looks at spring 2007 book tagging activity in LibraryThing.com.

Literature on social tagging [2], [9], and on the folksonomy [1], [21], [16] and [6] is burgeoning. Other recent studies consider social tagging for summary generation [3] and for information retrieval [10], [13], [23], [24], [26] and [27]. However, analyses of social tagging conducted thus far present little quantitative evidence as to strengths and weaknesses [21]. This study considers whether there is a difference between controlled and uncontrolled (enthusiast-generated) vocabulary in information retrieval judgments, and offers a method of evaluating user tag preference and the relative strength of social tag vs. LCSH string retrieval performance. The experiments produce quantitative evidence that reveal some of the risks and rewards that might follow from integrating social tagging into an academic library catalog.

II. COLLABORATIVE TAGGING (OR SOCIAL TAGGING)

A. Applying tag terms

Tag terms in a library environment are added by subject catalogers or indexers. The need for subject descriptors in a catalog record appears in the formula for a bibliographic record established early in the 20th century by Charles Cutter. It was Cutter's original premise that catalog users wish to find all books a library has by a given author, on a given subject, in a given kind of literature and to assist in the choice of a book by edition or literary or topical character [5]. Name and title control in the library community is regulated by the *Anglo-American Cataloging Rules*, and subject is regulated by the Library of Congress Subject Cataloging Manual for Subject Headings in the *Library of Congress Subject Headings*.

Catalogers are skilled in choosing pre-fabricated strings of terms from an authoritative term list. Controlled vocabulary lists help ensure all items on a given subject appear under a given controlled vocabulary term, such as to improve what is known as recall.

Compare this to the rules and habits of collaborative tagging, also known as social tagging. Here, anyone can make up any term for any item. The following directions from

LibraryThing.com explain how to assign tags (from http://www.librarything.com/concepts.php)

> "Anything can be a tag—just type words or phrases, separated by commas. Thus one person will tag the *The DaVinci Code* "novels" while another tags it "trashy, religion, mary," and still another only "summer home."

So rather than one person—the cataloger or indexer in the library environment—assigning terms, many people might assign tags to the same item. The problem is that without a controlled vocabulary list from which to select tag terms, these people likely will come up with synonyms. The work of Furnas, Landauer, Gomez and Dumais showed that the likelihood that many will assign the same name to the object is less than one in five [8]. But the same tag might be assigned multiple times to a single item. It has been found that people who are in contact are more like to re-use another's tag term [14], and the way a system is designed also influences taggers' choice. The problem of assigning terms is compounded when people use parallel spellings. Walker calls tagging "feral hypertext" for being wild and out of control [25].

The sum of recurring social tags is called a folksonomy. The folksonomy reflects the voice of a large number of people, while the application of the tags allows taggers their individualism. Also, these tags do not involve data loss by compressing items in pre-set categories, but instead filter into categories afterward [20]. Classifications by folksonomy are arranged with facets rather than hierarchically, as is the Library of Congress Subject Headings (LCSH).

Many hobbyists may tag the same resource, in what Tonkin calls "broad" tagging service [23]. By contrast, professional cataloging departments in academic libraries tend to be understaffed [4], and items are tagged by one person with only a few terms taken from a controlled vocabulary list. Tonkin calls this a "narrow" tagging service [23].

Enthusiasts tend to tag items they like or dislike, and their attentions combined creates a tag cloud. This leaves some items un-tagged. The reason Sen provides is that taggers, and more frequently non-taggers, cannot think of a tag term [18].

B. Information retrieval, controlled and uncontrolled vocabulary

Information retrieval studies often concern recall and precision. Recall could be defined as the proportion of relevant items retrieved from all items available, whereas precision is the proportion of relevant items retrieved out of what was retrieved. The purpose of controlled vocabulary for subject terms ultimately is to improve both recall and precision [12]. Many studies have been done on precision; fewer on recall. This might be in part because low precision tends to be troublesome and obvious to searchers, whereas searchers do not invariably recognize when system recall is low and records are missing from the result set.

It is the cross referencing of synonyms and alternative spellings and use of overarching terms instead of similar terms that prevents scattering and improves recall. Such are built into controlled vocabularies such as LCSH. LibraryThing.com offers a feature to link authors, works or tags called *combining*, but observation of retrieved results suggests this feature is used rarely.

Term for term, controlled vocabulary aims to be precise due to the care taken in selecting authoritative terms. In the social tagging environment, precision is self-adjusted to some extent by that fact that people choose what to tag. That is, they often choose to tag what they know or have read, and so should tag accurately. The LibraryThing software compensates for precision by using font size and boldness to emphasize tags chosen by more people (which presumably are more precise). The number of people tagging makes inconsistent tags drop out statistically [17].

In a study of participants' use of social tags to find movies [18], only 19% of participants felt that tagging actually helped them find what they were looking for. Participants commented that they found the tags useful for purposes such as helping to express their own opinions, providing information about items, helping to organize that information, and helping them decide whether to consult an item.

III. Experiments and results

A. Background to the experiments: Data Collection and Limitations

The social tags were taken from LibraryThing.com and LCSH strings were taken mostly from the Library of Congress online catalog. All three experiments used instruments created for the purpose. The books selected for the experiments were selected on topics of general interest selected randomly. Within these topics, particular books were chosen for having tag clouds. Not every item in the LibraryThing.com catalog has tags, and some of those which do have only a few tags. The fewer the tags, the more likely it is that the tag terms are inaccurate or idiosyncratic. Therefore, the survey instrument if anything favors tag clouds over strings. This makes some of the experimental findings all the more surprising.

The first experiment uses a standard 7-point Likert scale to rate strength of opinion on catalog recall. Experiment 2 uses four basic questions with the same answers in different forms (three books, or three strings for those books, or three clouds for those books). Experiment 3 combines social and LCSH tags for the same book and asks participants to choose which tags are most descriptive.

The questions apply to online catalogs, but are not asked in the context of online catalogs specifically. This is because indexing works differently in standard academic library catalog software and in LibraryThing, and the goal is to compare not the varying retrieval algorithms, but rather to compare the tags themselves. In that the LibraryThing and Library of Congress catalogs have different purposes, taking the books, tags, and questions into a single context helps to balance the indexing found in the two catalogs.

B. Experiment 1: Recall

Do participants care whether all items relevant to a query are retrieved in an online catalog search?

User database queries will likely not retrieve every appropriate social-tagged match because the social tags are not entered consistently and recall is low. The question is: does it matter to catalog searchers?

People who were in mid-search on an academic library catalog were asked to pause to answer a short questionnaire on their opinion of catalog recall. The word "recall" was not used in the instrument because its definition in the information retrieval context is not known commonly. The question was "How do you feel about whether the catalog search retrieves every single record/item that is relevant? (Please circle a number)." The scale ran from 1 – 7, with 1 unimportant, 4 neutral and 7 highly important.

Of the 39 people surveyed, the average level of importance assigned to getting comprehensive search research is 5.8 with a standard deviation of 1.2.

Participants sampled showed a difference in opinion as to importance of search recall depending on whether their search was for research or for personal interest. Findings suggest that those who search the catalog for research are more likely to seek comprehensive results than those who search for pleasure. But even those who search for pleasure on average care somewhat that the catalog results displayed be comprehensive.

Among participants, gender was not predictive of opinion as to importance of recall. Age range seems to have some bearing on whether searchers want results to be comprehensive. The range of responses was greatest among younger searchers, but the average response was that it is important (6 on Likert scale of 7) for results to be comprehensive. In the given sample, it was most important to those in their 40s at mid-career (7 on Likert scale of 7), and less important to searchers age 61 and older that results are comprehensive (5 on Likert scale of 7).

C. Experiment 2: Tag format preference and performance

1) Background to the shared instrument and the second experiment

Adding headings increases precision [11], [8]. Lancaster suggested that it is the number of access points that could increase retrievability as much if not more than the type of vocabulary [11]. So one might surmise that the cloud format with its large number of terms would be preferred to the string format. But this was not borne out by the experiment.

The survey instrument has three parts, the parts are delivered sequentially, and the same 4 questions recur in each part. The first part concerned preference for tag format, the second and third parts concerned how well the tag format "performed" in predicting choice of actual book.

Topics (bird flight, various ethnicities of Latin America, the archaeology of ancient Greece and global warming) were selected in an effort to appeal to a wide range of participants. The questions aimed to simulate questions of quotidian interest, rather than recondite research. In the survey, questions recur on three different instruments which participants were handed and completed sequentially, so that participants likely were unaware of or at least, not bothered, by the repetition.

The experiment is limited by whether the research questions are realistic. It is limited also in whether the clouds are more effective in describing fiction than nonfiction, even though the research questions used were nonfiction.

To lessen the possibility that the answer to each question was determined by a particular tag cloud or string rather than the cloud or string form in general, participants were asked to perform each task four times for four different cloud and string sets. The unit of analysis, then, is the relevance judgment that inspired choice of cloud, string, or book. Twenty-three people participated, so there were 23 x 4, or 92 observations.

2) Preference

Do participants prefer tag clouds or strings?

The test was to determine whether cloud or string was preferred to answer the question. Participants were not told why some cloud words were in bolder, larger font. Neither were they told the meaning of the dashes in the hierarchical string.

It was found, as anticipated, that the choice of cloud or string depends more upon the particular cloud or string than the appeal of the cloud or string format alone. 18 people preferred cloud or string descriptions of the book, depending upon the cloud or string. 5 people selected one format only regardless of the particular cloud or string (of these, 3 preferred string exclusively, 2 preferred clouds).

Patterns were evident in the selection of particular clouds and strings. For three of the four books selected (birds, archaeology, global warming), participants preferred strings to clouds by about 3 to 1, and for one book (Latin America), participants preferred clouds by almost the same 3 to 1 margin.

The quality of the string or cloud might have something to do with the fact that Library of Congress catalogers often have subject knowledge in an area. Or it might have something to do with the strength of the social tags in the area of modern culture.

Altogether there were 92 choices of cloud or string format, but four were omitted by participants, leaving 88 relevance judgments. Strings were preferred predominantly for 55 of those choices, or 63% of the time.

What reason did participants give for their choice of cloud or string format? Most participants made their selections because they felt the words comprising the set they chose were more descriptive than the words in the other set. Also the number of words in the cloud or string often could be influential in the decision. Only a few participants were influenced by the random shape of the cloud or the formal shape of the string. In addition to these factors, participants were given an option following each question to write a brief explanation of what influenced their preference for cloud or string. Strings prompted many more write-in explanations than the clouds. Write-in reasons for selecting the strings included "concise," "neater," "organized," "comprehensible," "specific,"

"presentable," "basic information," "categorization," and "relevant". In some cases, the selection of the string was prompted by the dislike of the cloud, which was called a "distraction," "length" (implying the cloud was too large) and "information overload". One person commented that the clouds "seem like spam". The few positive write-ins for clouds pointed to the "meaning" of the words, and the fact that the cloud is "eye catching".

3) Performance of tags for book selection

Are clouds or strings more useful for making judging relevance?

In terms of performance, we want to learn whether the clouds and strings "perform" adequately in describing what the book was about, and whether one format out-performs or whether the two are equivalent.

The instrument repeats the four general questions about birds, Latin America, archaeology and global warming. To answer each question, the participant was given a choice of three clouds (each corresponding to a book that should answer the question), and chooses that cloud set that answers the question best. After this he was given a choice of three strings (each corresponding to a book that should answer the question), and chooses the string that answers the question best. Finally, he is presented three books to answer the question (the books, unknown to him, correspond to the same cloud and string sets he was presented earlier).

This measure of performance is imperfect in that people do not select books on the basis of content only, and so the strength of moderating factors in the decision such as publication date and number of illustrations are considered as well.

When a participant selected the same cloud or string set as book to answer the question, that cloud or string set is said to "perform." According to this definition, there is no measure of performing to a greater or lesser degree; the format either performs by corresponding to the book the participant selected, or it does not perform and does not correspond to the book selected.

The concept of how a cloud or string "performs" was limited and perhaps crippled by the fact that it assumed that cloud and string terms are assigned based primarily on content, and that books are selected to answer questions based primarily on content. Both these assumptions are weak. LibraryThing explains to users that "anything" can be a cloud term. (This not the case for strings, where subject catalogers assign LCSH descriptors primarily on the basis of content.) In addition, participants noted the relative importance of factors influencing their decision as to choice of book, factors which are not generally reflected in index terms. Such factors included physical condition, date of publication, number of illustrations and length, layout, the presence of maps or index, the quality of illustrations or a particular publisher.

The number of people who participated was 23 (N=23), and with four possible judgments per topic question, the number of possible observations should be 92 (n=92). However, the some questions were omitted. The questions were given in the same order every time, with the global warming question last, so it was mostly the global warming responses omitted for those who did not finish the study. (6 global warming responses, 3 bird responses, 1 archaeology response and 1 Latin America response were omitted. Therefore, with 11 responses omitted, the number of possible observations was reduced to 81 (n=81).

Insofar as these figures can be said to represent "performance" of cloud and string, strings did better than clouds in three of four tasks, and overall, strings "performed" 38% of the time, while clouds "performed" 25% of the time.

Findings suggest that this measure of performance might not be suitable, and that the test was not reliable: that is, it did not measure what it was supposed to measure. Correspondence between selected cloud or string to selected book was low.

D. Experiment 3: Words of the folksonomy or LCSH

Do participants find collaborative or controlled vocabulary terms more useful for judging relevance?

The instrument was created based on the random selection of 9 books from the catalogs, the requirement being that the book had to have been given headings both in LibraryThing and by catalogers from the Library of Congress. The distribution of tags per book is not uniform. That is, some books had more Library of Congress headings than social tags. Summaries that included few if any descriptor terms were created for each book. Book titles were shorted in some cases if the title contained descriptor terms. For each of the 9 books following title and summary, social tags and Library of Congress subject headings were intermixed randomly. Terms that overlapped on the LibraryThing and Library of Congress indexing were removed, and to lessen noise, some LibraryThing tags with little overt meaning were excluded. There were 117 tags in total, of which 32 were from LC. That makes 73% social tags and 27% LC subject headings. Participants were asked to select the three terms most descriptive of content.

Data was collected as follows. N=36 and because each person had 9 (books) x 3 (tag choices per book), it should be that n=972. In fact, eight questions were unfinished, and a few others partially finished with only two instead of three terms selected per book, reducing n such that n=940.

The final tally was: Collaborative tags selected = 547; LCSH terms selected = 393. This means that 42% of the possible points chosen were from LCSH, even though LCSH comprised only 27% of the terms total. Thus, if participants had chosen terms at random, they would have chosen LCSH 27% of the time. But they chose LCSH terms 42% of the time, considerably more frequently (42/27 or 1.6 times as frequently).

IV. DISCUSSION

Findings of experiment 2 showed that participants preferred what some referred to as the neat, organized, comprehensible string format to what some referred to as a distracting, large

cloud. Findings that neither string nor cloud format are a good basis for judging the book probably shows less about the supposed performance of either string or cloud than that the method to weigh performance was flawed. Findings of experiment 3 showed that, word for word, the controlled vocabulary of LCSH was preferred to the freely chosen collaborative tags.

Even though catalog users might prefer a collaborative tag format other than the cloud, and might find collaborative tags not particularly effective in describing item content, the use of collaborative tagging in catalogs could be justified if users were equivocal about recall (which is weak in non-controlled vocabularies such as collaborative tags). But findings of experiment 1 showed that searchers of academic catalogs felt that high recall was important, whether they were doing scholarly or non-scholarly work.

In sum, preliminary findings based on the small sample in this study suggest that social tags clouds in their current form, while better than no tags at all, are inferior to the Library of Congress Subject Headings strings for information retrieval. This is not to say that we could not harness value in collaborative tags constructively. Collaboratively-supplied words have been used for summaries [3], and the software helps create value.

These findings might be explained in part by the fact that the significance of the purpose of the randomly-space cloud words that differ in font size and thickness is not widely understood or appreciated. System designers might consider writing a program that arranges collaboratively tags hierarchically or after the better-liked strings of the LCSH model. Even the similarities and overlaps in terms could be filtered by software.

How could we proceed based on these findings? Could the Library of Congress Subject Headings or other controlled vocabularies be employed more widely for information retrieval? Granted, there are costs in (training and) employing catalogers. But the headings are used over and over, and benefits accumulate over time as more searches are made. In the long run, that increase in usefulness adds up and so the increased cost in applying controlled vocabulary might be economical, not only for a library, but for a commercial database as well. A tentative conclusion might be that professional indexing of entire web sites might outperform tag cloud indexing.

Findings not withstanding, taggers might step in where professional indexing costs are prohibitive. Would automated generation of tag clouds be any improvement over present keyword indexing? Future researchers must find out.

V. Conclusion

This study contributes a method for quantitative analysis of collaborative tagging. It compares the information retrieval value of the cloud format tags and the tag words themselves as found in the LibraryThing catalog (with low information recall properties of uncontrolled vocabulary) to the hierarchical string format and words themselves in the Library of Congress Subject Headings in the Library of Congress online catalog. Results are that, in terms of format, the staid strings of LCSH are preferred to the patchy clouds of social tagging, and that the words within the strings were better descriptors than the words within the clouds. Results are also that, whether searchers are working toward research or personal ends, high recall matters. Another investigation with larger sample size should be conducted to confirm results.

In light of this preliminary study, recommendations are that controlled vocabulary such as LCSH be used instead of social tagging when information retrieval from the catalog is important. Or to improve information retrieval, system designer might improve the way existing tags are revealed to those contemplating what new tags to enter, a way for synonyms to be linked to improve recall, and a way to display tag terms to users other than with word size and font such that word salience is more understandable. Finally, collaborative tags could be harnesses for purposes beyond information retrieval; this study should encourage others to conduct further investigations.

Acknowledgment

I would like to express my appreciation to Michael Lesk and to Ned Wall for their assistance and support.

References

[1] X. Arch, "Creating the academic library folksonomy," *College & Research Library News,* vol. 68(2), 2007, pp. 80-81.

[2] M. E. Bates, "Tag—You're It!" Online, vol. 30(1), 2006, p. 64.

[3] W. Boydell and B. Smyth, "From social bookmarking to social summarization: An experiment in community-based summary generation," *Proceedings of the 12th International Conference on Intelligent User Interfaces, Jaunary 28-31, 2007, Honolulu Hawaii*, 42-51.

[4] J. Byrd et. al., "A white paper on the future of cataloging at Indiana University." 2006. Retrieved February 8, 2007 from http://www.iub.edu/~libtserv/pub/Future_of_Cataloging_White_Paper.pdf

[5] C. A. Cutter, "Rules for a dictionary catalog," In M. Carpenter and E. Svenonius, Eds. *Foundations of Cataloging: A Sourcebook*, Littleton, CO: Libraries Unlimited, 1985 pp.65-71.

[6] "Folksonomy: A game of high tech (and high stakes) tag." *EConten*t, vol. 29(3), 2006, pp. 38-43.

[7] G. W. Furnas et. al., "Why do tagging systems work?" *Conference on Human Factors in Computing Systems, April 22-27, 2007, Montréal Québec, Canada,* pp. 36-39.

[8] G. W. Furnas, T. K. Landauer, L. M. Gomez, and S. T. Dumais, " The vocabulary problem in human-system communication." *Communications of the ACM,* vol. 30(11), 1987, pp. 964-971.

[9] S. A. Golder and B. A. Huberman, "Usage patterns of collaborative tagging systems," *Journal of Information Science*, vol. 32(2), 2006, pp.198-208.

[10] Hotho et. al., "Information retrieval in folksonomies: Search and ranking," 2006. Retrieved March 4, 2007 from http://www.kde.cs.uni-kassel.de/stumme/papers/2006/hotho2006information.pdf

[11] F. W. Lancaster, *Indexing and Abstracting in Theory and Practice*. 2nd ed. Champaign, Ill., Graduate School of Library and Information Science, 1998.

[12] F. Leise, "Metadata and content management systems: An introduction for indexers. *The Indexer*, vol. 24(2), 2004, pp.71-74.

[13] G. Macgregor and E. McCulloch, "Collaborative tagging as a knowledge organization and resource discovery tool. *Library Review*, vol. 55(5), 2006, pp. 291-300.

[14] C. Marlow, M. Naaman, D. Boyd, and M. Davis, " HT06, tagging paper, taxonomy, flickr. Academic article, to read," *Proceedings of the seventeenth conference on hypertext and hypermedia, August 22-26, 2006, Odense, Denmark,* pp.31-40.

[15] T. A. Peters and M. Kurth, "Controlled and uncontrolled vocabulary subject searching in an academic library online catalog," *Information Technology and Libraries*, vol. 10, 1991, pp. 201-211.

[16] E. Peterson, "Beneath the metadata: Some philosophical problems with folksonomy," *D-Lib Magazine* vol. 12(11), 2006, Retrieved February 19, 2007 from http://www.dlib.org/dlib/november06/peterson/11peterson.html

[17] M. L. Rethlefsen, "Chief Thingamabrarian," *Library Journal*, vol. 132(1), 2007, pp. 40-42.

[18] S. Sen, S. K. Lam, A. M. Rashid, D. Cosley, D. Frankowski, J. Osterhouse, F. M. Harper, J. Riedl, "Social tagging and recommending: Tagging, communities, vocabulary, evolution," *Proceedings of the 20th anniversary conference on computer supported cooperative work, November 4-8, 2006, Banff, Alberta, Canada* (pp 181-190). New York: ACM Press.

[19] W. M. Shaw Jr., "Controlled and uncontrolled subject descriptions in the CF database: A comparison of optimal cluster-based retrieval results," *Information Processing & Management*, vol. 29 (6), 1993, pp.751-763.

[20] C. Shirky, "Clay Shirky's writings about the Internet: Economics and culture, media and community," 2005. Retrieved March 26, 2007 from http://www.shirky.com/writings/ontology_overrated.htm

[21] L. F. Spiteri, "The use of folksonomies in public library catalogues," *The Serials Librarian*, vol. 51 (2), 2006, pp.75-89.

[22] D. Sturz, "Communal categorization: The folksonomy," 2004. Retrieved March 4, 2007 from http://www.davidsturtz.com/drexel/622/sturtz-folksonomy.pdf

[23] E. Tonkin, "Folksonomies: The fall and rise of plain-text tagging," 2006. *Ariadne,* vol. 47 Retrieved February 19, 2007 from http://www.ariadne.ac.uk/issue47/tonkin

[24] J. Trant, "Exploring the potential for social tagging and folksonomy in art museums: Proof of concept," *New Review of Hypermedia and Multimedia,* vol. 12 (1), June 2006, pp. 83-105.

[25] J. Walker, "Feral hypertext: When hypertext literature escapes control," *Proceedings of the Sixteenth ACM Conference on Hypertext and Hypermedia, September 6-9, 2005, Salzburg, Austria,* pp. 46-53.

[26] H. Wu, M. Zabair and K. Maly, "Harvesting social knowledge from folksonomies," *Proceedings of the seventeenth conference on hypertext and hypermedia, August 22-26, 2006, Odense, Denmark*, pp. 111-114.

[27] Y. Yanbe, A. Jatowt, S. Nakamura and K. Tanaka, "Can social bookmarking enhance search in the web?" *Proceedings of the 2007 International Conference on Digital Libraries, June 18-23, 2007, Vancouver, British Columbia, Canada*, pp. 107-116.