

Does Catalog Tagging Aid Retrieval for ESL Users?

A Research Proposal for an Experimental Study

XXX XXX

San Jose State University

School of Library and Information Science

Table of Contents:

1. Introduction	3
2. Literature Review.....	3
3. Definition of Research Problem.....	4
4. Methodology.....	5
5. Execution Plan.....	10
6. Conclusion.....	10
7. References.....	12
8. Appendix A: Informed Consent Form	16
9. Appendix B: Demographic Information Form.....	18
10. Appendix C: Measuring Instrument.....	21

Introduction

Tagging and folksonomies have made significant changes in how we think about information retrieval, offering a vision of a world where people en masse control the organization and retrieval of data. Yet the effect of tagging on the searching efforts of ESL users has not been previously examined. This research proposal outlines an experimental study to determine whether a catalog enhanced with user-generated tags would aid ESL students at a community college in locating sources for a research paper in a composition course. This seems a worthwhile user group to investigate, given the increasing linguistic diversity of American society and the globalizing effect of the Internet.

The research project is predicated on the assumption that a catalog of suitable user-tagged articles exists; most likely, this will entail a previous study in which users similar to the subject population (community college students, both ESL and native speakers) generate tags for articles likely to be retrieved in a given search (in this case, a research paper for a composition class).

This paper will review the literature on the topic of tagging, define the research problem, describe the project's methodology, outline the plan for executing the project, and detail the contributions of the project and its implications for the field.

Literature Review

Folksonomies and tagging began as part of what is broadly known as Web 2.0, which revolutionized information organization and retrieval and posed a challenge to traditional knowledge organization systems. Tagging originated as a way to share online bookmarks (Steele, 2009, p. 69). Thomas Vander Wal coined the term "folksonomy" to describe the collection of

terminology created by users and to distinguish it from traditional taxonomies and ontologies (Trant, 2008, p. 1). Tags had many detractors, who noted that tags did not control for synonyms or polysemy (Steele, 2009, p. 71), were often ambiguous and did not show relationships between concepts (Daly, 2009, p. 5), had too much granularity (De Meo, Quattrone & Ursino, 2009, p. 512), and disproportionately represented techno-savvy users (Rolla, 2009, p. 176). However, people were also quick to point out the advantages of tagging, including a reflection of users' natural vocabulary (Daly, 2009, p. 5), the increase in the number of access points (Steele, 2009, p. 77), and the rapid adaptation of new terms and representation of minority terminology (Trant, 2008, p. 8). In terms of information retrieval, Sinclair and Cardew-Hall found that a tag cloud was useful for browsing because it reduced the cognitive load on the user, but it was less useful for specific searches (2008, p. 27-28). The current trend is toward examining the convergence of controlled vocabulary systems and natural language folksonomies (Leong, 2010, p. 39).

However, there is a significant gap in the literature in terms of addressing linguistic diversity. Only one article addresses translation issues (Noh et al., 2009), and no other articles consider non-native English speakers or ESL users at all. Yet if, as the research suggests, tagging uses more intuitive vocabulary, reduces the cognitive load, and provides additional points of entry, could it have particular relevance for ESL users?

Definition of Research Problem

The first step is to put the research questions in terms of constructs: For ESL users, does a tag cloud interface provide superior retrieval over a traditional keyword/subject term catalog interface in terms of ease of use, precision, and recall?

The questions can then be posed as a series of conceptual hypotheses:

Ease of Use

H₀: The tag cloud and traditional interface are equally easy to use.

H_a: The tag cloud interface is easier to use.

Recall

H₀: Both interfaces have equal recall.

H_a: The tag cloud interface has superior recall.

Precision

H₀: The traditional interface retrieves more precise results (or no difference).

H_a: The tag cloud interface retrieves more precise results.

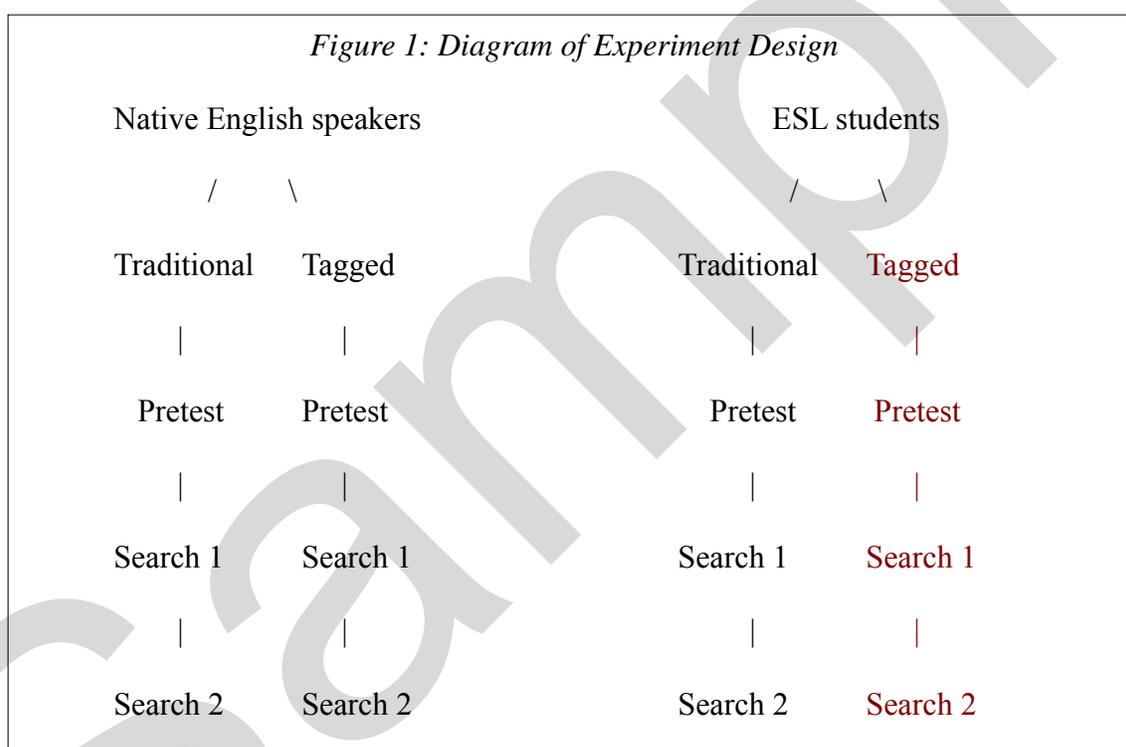
The alternative hypotheses will be accepted if there is a statistically significant difference at the 95% level ($p < 0.05$). It seems probable that there will be a statistically significant improvement in ease of use and recall, but not necessarily precision.

Methodology

Research Design

The study is based on a classical experiment design. In order to isolate whether a tagged catalog was useful for ESL students, one would have to control for both language background and catalog design. Therefore, students would be split into two groups: native English-speaking students and ESL students. Students from each group would then be randomly assigned to search using a traditional catalog or a catalog enhanced with user-generated tags. An initial pretest would be conducted where all groups use the traditional catalog interface to search for sources for a research paper for their composition class. This pretest would establish a baseline and

ensure that all groups are comparable and free of any significant differences. In the post-test, the researcher would divide the groups according to the interface they were testing (traditional or tag-enhanced), have assistants instruct students in the use of the interface to be tested, and have each group conduct two rounds of searching for additional sources. Having two rounds of searching would test for any learning effect and allow for both groups to be on equal footing if one interface has a steeper learning curve.



The experiment has high external validity. Although the search process itself is structured and artificial, the search itself is part of a real assignment, which motivates users to take the search seriously and to complete it as well as they can. The experiment would also have high internal validity. Students would be assigned to their group through simple random sampling. They would be trained in using each interface by an assistant who was unfamiliar with the

hypothesis and therefore could not communicate researcher expectation. Students would also be instructed not to talk about the experiment with each other.

Research Population and Sampling

The research population is of native English speaking and ESL students in community colleges. However, the practical population for sampling is of ESL students at one particular community college taking a composition class. This decision of limiting practical population was for multiple reasons, mostly because community college is a good place to find a high percentage of non-native English speakers. Also, community college students represent a wide range of ages, experiences, and linguistic and computer proficiencies, which would make the findings more generalizable to the library user population as a whole. The composition class was chosen as a useful subset because it is a prerequisite for all students, the research paper would require students to conduct focused searches, and a research paper for a composition class would generate a wider range of topics than papers for other classes would. This allows for a broader data set.

Every student taking a composition class during the designated semester would be eligible to be selected for the experiment. The students would be divided into groups of ESL students and native English speakers and assigned random numbers, which would be entered into a database and selected at random to be placed in one of the four groups (English/traditional, English/tagging, ESL/traditional, ESL/tagging). The selected participants would be given an informed consent form to sign (see Appendix A). Those who chose not to participate would be replaced by other randomly selected participants. (If there was a high turnover rate, researchers would consider using volunteers.) The experiment would require a total of 64 students: 40 ESL

participants (20 for each group) and 24 native English-speaking participants (12 for each group).

Interface Designs

The fact is that interface design does have an effect on how users search (Knievel, Wakimoto & Holladay, 2009, p. 458), which means the design would have to be carefully considered. The traditional interface need provide no more options than a standard library catalog: a search bar that offers keyword, title, and subject term searching (although participants are unlikely to use the latter two). The enhanced tagging interface may be a simplified version of the one used by Matthews et al. in the Intute study (2010, p. 453), containing both a global tag cloud and a tag search box. However, having essentially two different tag search options may complicate the data, and so one search option may ultimately be selected depending on the perceived needs of the users and the study.

Data Collection

Before the pretest, each student would be given a demographic information form to collect information about the student's age, sex, race/ethnicity, native language, English proficiency, and computer proficiency (see Appendix B). Ideally, the data for the last two criteria would be objectively confirmed by placement test scores and a brief computer test, but if no testing instrument is readily available or feasible within the time frame, the self-reported data alone may be deemed sufficient. The data would be kept private by identifying participants on the information forms only by their randomly assigned user number. Only the head researcher would have access to the information linking the participant to the user number.

During the two rounds of testing, a query log toolbar will record search terms used,

sources retrieved, and time spent by subjects examining each source. Each interface will also be designed so that a series of check boxes appears next to each source: Relevant/Useful—Yes, Maybe, No. The participants would be asked to decide whether the source is relevant or useful for their paper and check the appropriate box. For the tag-enhanced interface, the query log toolbar would also note the number of tags for each article retrieved in order to determine whether there is a certain tag density necessary to maximize retrieval.

After the completion of the second round of tests, participants would be given a post-test survey to determine their satisfaction with the ease of searching, overall recall, and search precision. The survey has a closed-ended component using a 5-point Likert scale and an open-ended component where participants are encouraged to elaborate on their experiences with the search interface (see Appendix C).

Data Analysis

The data analysis process would be complex. One step would involve cross-tabulating English proficiency, computer proficiency, and ease of use data to discover any relationships between those factors. Recall would be measured by examining the query log data and determining the number of sources retrieved for each query. Precision would be calculated by dividing the number of sources deemed relevant for a query by the total number of sources retrieved. These data would be compared to the user satisfaction ratings for recall and precision. Researchers would also create a frequency distribution table for number of tags per retrieved article; if a significant number of sources have at least a certain number of tags, further tests would have to be run to examine the effect of tag density on retrieval. A mean comparison would be run to determine whether there were any statistically significant ($p < 0.05$) differences

between the native English speakers and the ESL students. Using the statistical hypothesis testing listed in the "Definition of Research Problem" section, inferential statistics would be run to determine whether there is any statistically significant ($p < 0.05$) variation in ease of use, precision, or recall between the traditional and the tagging interfaces.

Execution Plan

The first step would be to locate a community college library to serve as a research site. There may be some incentive for community college libraries to consider participating, since they are currently underrepresented in the LIS literature and rarely receive the opportunity to participate in large research projects. The community college itself would need to be in a diverse area and have a high percentage of non-native English speakers in attendance.

The experiment would require a developer to create the tag-enhanced interface, taggers to tag articles (if this has not already been done as part of a previous study), a research assistant to aid in coding and analysis, an assistant or library worker to instruct participants in the use of the interfaces, access to a computer lab, a printer, and a copier. Further costs include a small payment of \$10 to each participant. Depending on what resources the library already had available, the cost could be anywhere from \$4,000 to \$10,000. The estimated timetable is two years for development, six months of which would be devoted to the execution of the experiment. The participants would be actively engaged in the experiment for one month of that time.

Conclusion

This experiment could contribute much to the profession's understanding of tagging as it relates to non-native English speakers. It examines the potential benefits of "next generation" catalogs to understudied user group, provides a foundation for further studies of catalog design for multilingual populations, and provides the impetus for further tagging studies to address multilingual and ESL issues. If the study finds statistically significant differences in retrieval effectiveness between the interfaces, then it is possible that catalogs can be made more friendly to multilingual populations without translating the entire catalog into other languages. It will also provide valuable quantitative information about precision and recall for traditional versus folksonomy-enhanced databases. Furthermore, it offers some quantitative information about the saturation point of tags and how many tags a particular item must have before it becomes findable, although that information is not the main point of the study.

A potential issue is the fact that users often have difficulty judging the relevance of documents that are not in their native language (Hansen & Karlgren, 2005, p. 635). Future studies about multilingual users may wish to triangulate self-described relevance data with a more objectively measurable method.

On the whole, the study would simple, measurable data concerning ESL students' use of catalogs and analyze whether tagging aids in retrieval in a specific case. However, there are a wide range of hybrid knowledge organization systems, and much work would remain to be done to determine the best retrieval system to meet the needs of this particular user group.

References

- Aharony, N. (2009). Librarians and information scientists in the blogosphere: An exploratory analysis. *Library & Information Science Research*, 31(3), 174-181.
doi:10.1016/j.lisr.2009.02.001.
- Daly, E., & Ballantyne, N. (2009). Ensuring the discoverability of digital images for social work education: An online "tagging" survey to test controlled vocabularies. *Webology*, 6(3), 1-16. Retrieved from Library, Information Science & Technology Abstracts with Full Text database.
- De Meo, P., Quattrone, G., & Ursino, D. (2009). Exploitation of semantic relationships and hierarchical data structures to support a user in his annotation and browsing activities in folksonomies. *Information Systems*, 34(6), 511-535. doi:10.1016/j.is.2009.02.004.
- Dotsika, F. (2009). Uniting formal and informal descriptive power: Reconciling ontologies with folksonomies. *International Journal of Information Management*, 29(5), 407-415.
doi:10.1016/j.ijinfomgt.2009.02.002.
- Echarte, F., Astrain, J., Córdoba, A., & Villadangos, J. (2008). Self-adaptation of ontologies to folksonomies in Semantic Web. *Proceedings of World Academy of Science: Engineering & Technology*, 43, 325-331. Retrieved from Library, Information Science & Technology Abstracts with Full Text database.
- Hansen, P., & Karlgren, J. (2005). Effects of foreign language and task scenario on relevance assessment. *Journal of Documentation*, 61(5), 623-639.
- Kniesel, J. E., Wakimoto, J. C., & Holladay, S. (2009). Does interface design influence catalog

- use? A case study. *College and Research Libraries*, 70(5): 446-458.
- Kwan, Y., & Chan, L. M. (2009). Linking folksonomy to Library of Congress subject headings: An exploratory study. *Journal of Documentation*, 65(6), 872-900.
- Lawson, K. G. (2009). Mining social tagging data for enhanced subject access for readers and researchers. *Journal of Academic Librarianship*, 35(6), 574-582.
- Leong, J. H. (2010). The convergence of metadata and bibliographic control? Trends and patterns in addressing the current issues and challenges of providing subject access. *Knowledge Organization*, 37(1), 29-42.
- Matthews, B., Jones, C., Puzoń, B., Moon, J., Tudhope, D., Golub, K., & Nielsen, M. L. (2010). An evaluation of enhancing social tagging with a knowledge organization system. *Aslib Proceedings: New Information Perspectives*, 62(4/5), 447-465. doi: 10.1108/00012531011074690
- McFadden, S., & Weidenbenner, J. V. (2010). Collaborative tagging: Traditional cataloging meets the "wisdom of crowds." *The Serials Librarian*, 58, 55-60. doi: 10.1080/03615261003623021
- Mendes, L., Quiñonez-Skinner, J., & Skaggs, D. (2009). Subjecting the catalog to tagging. *Library Hi Tech*, 27(1), 30-41. Retrieved from Library, Information Science & Technology Abstracts with Full Text database.
- Merčun, T., & Žumer, M. (2008). New generation of catalogues for the new generation of users: A comparison of six library catalogues. *Program: Electronic Library & Information Systems*, 42(3), 243-261. doi:10.1108/00330330810892668.
- Morrison, P. J. (2008). Tagging and searching: Search retrieval effectiveness of folksonomies on

the World Wide Web. *Information Processing and Management*, 44, 1562-1579.

Retrieved from Library, Information Science & Technology Abstracts with Full Text database.

Moulaison, H. L. (2008). Social tagging in the Web 2.0 environment: Author vs. user tagging. *Journal of Library Metadata*, 8(2), 101-111. doi: 10.1080/10911360802087325

Noh, T.-G., Park, S.-B., Yoon, H.-G., Lee, S.-J., & Park, S.-Y. (2009). An automatic translation of tags for multimedia contents using folksonomy networks. *SIGIR Forum*, 492-499.

Retrieved from Library, Information Science & Technology Abstracts with Full Text database.

Pera, M. S., Lund, W., & Ng, Y.-K. (2009). A sophisticated library search strategy using folksonomies and similarity matching. *Journal of the American Society for Information Science & Technology*, 60(7), 1392-1406. doi:10.1002/asi.21072.

Rolla, P. J. (2009). User tags versus subject headings: Can user-supplied data improve subject access to library collections? *Library Resources & Technical Services*, 53(3), 174-184.

Retrieved from Library, Information Science & Technology Abstracts with Full Text database.

Schwartz, C. (2008). Thesauri and facets and tags, oh my! A look at three decades in subject analysis. *Library Trends*, 56(4), 830-842. Retrieved from Library, Information Science & Technology Abstracts with Full Text database.

Sinclair, J., & Cardew-Hall, M. (2008). The folksonomy tag cloud: When is it useful? *Journal of Information Science*, 34(1), 15-29. Retrieved from Library, Information Science & Technology Abstracts with Full Text database.

Steele, T. (2009). The new cooperative cataloging. *Library Hi Tech*, 27(1), 68-77. doi:

10.1108/07378830910942928

Thomas, M., Caudle, D. M., & Schmitz, C. M. (2009). To tag or not to tag? *Library Hi Tech*,

27(3), 411-434. doi: 10.1108/07378830910988540

Trant, J. (2009). Studying social tagging and folksonomy: A review and framework. *JODI:*

Journal of Digital Information, 10(1), 4. Retrieved from Library, Information Science & Technology Abstracts with Full Text database.

Tsui, E., Wang, W., Cheung, C., & Lau, A. (2010). A concept-relationship acquisition and

inference approach for hierarchical taxonomy construction from tags. *Information*

Processing and Management, 46, 44-57. Retrieved from Library, Information Science & Technology Abstracts with Full Text database.

Appendix A: Informed Consent Form

December 13, 2010

Dear Participant:

You have been randomly chosen to be a part of a study about what types of online library catalogs are best for helping people find information. This is a very important study that will help scientists and librarians understand how different groups of people look for information so they can organize it in ways that will help people find it easily. The information we collect from you during the study may have a direct effect on what kind of library catalog students at this school and others like it will have access to in the future.

First, we will have you fill out a short survey about your English skills, computer skills, and background. All of the information will be kept completely private. We will give you a user number at the beginning of the study, and you put it on any forms we give you instead of your name. Only I will know which user number goes to what name, and I will keep that information secret. No information that could be used to identify you will be published.

The next step is meeting in the computer lab once a week for 3 weeks and using the catalog system we show you to search for sources to use in your research paper. Each session will last about 30-45 minutes. An assistant will be there to teach you how to use the system. A program on the computer will record what things you search for. This may make you uncomfortable, but you will not be recorded, and you will use your user number instead of your name or student ID to log in.

After you finish the last search session, we will ask you to fill out a two-page survey describing your experiences with the catalog system. Most of the questions only ask you to circle an answer, but for the last four, you will be asked to write a little bit about your experiences using the search system. If you would prefer not to write, you will be allowed to tell your answers to a research assistant, who will write them down for you.

After you have completed the sessions and filled out the final survey, you will be paid \$10 for your participation. Your teacher may also choose to offer extra credit for helping with this study.

You have the right to choose not to be a part of this study. No one is allowed to force you to participate.

If you choose to be a part of the study, but later change your mind, you will be allowed to leave. No one can force you to stay.

You have the right to see the results of the study when they are published, if you choose.

Please fill out the section below, sign it, and return it to me or one of my research assistants. If you have any questions or concerns about this study, please contact me.

Sincerely,

XXX XXX

xxx@gmail.com

-
- I have read this consent form.
 - I understand the purpose of the research.
 - I understand my rights and responsibilities as a part of this study.

_____ I wish to participate in this study.

_____ I do not wish to participate in this study.

Your name: _____

Your signature: _____

Today's date: _____

Appendix B: Demographic Information Form

User Number: _____

Age: _____

Sex: Male Female

How would you describe your race/ethnicity? (Circle one or write in an answer for "other.")

White/Caucasian Black/African-American Asian/Pacific Islander Native
American/American Indian/Alaska Native Hispanic/Latino Other: _____

What language(s) do you speak fluently besides English? _____

(If you filled in something for the last question, please fill out Part 1 AND Part 2. If you left this question blank, please skip Part 1 and fill out Part 2 ONLY.)

Part 1: English Proficiency

Please circle the answer that describes your situation best.

1. My first language was a language other than English.

Yes No I learned English and another language at the same time

2. When I first started learning English, I was:

0-5 years old 6-10 years old 11-15 years old 16-20 years old

21-25 years old More than 25 years old

2. I do most of my thinking in a language that is not English.

Strongly agree Agree Don't agree or disagree Disagree Strongly disagree

3. I have trouble understanding spoken English.

Strongly agree Agree Don't agree or disagree Disagree Strongly disagree

4. I have trouble speaking English.

Strongly agree Agree Don't agree or disagree Disagree Strongly disagree

5. I have trouble reading English.

Strongly agree Agree Don't agree or disagree Disagree Strongly disagree

6. I have trouble writing in English.

Strongly agree Agree Don't agree or disagree Disagree Strongly disagree

Part 2: Computer Proficiency

Please circle the answer that describes your situation best.

1. I use computers:

Every day A few times a week Once a week A few times a month Once a month
Less than once a month

2. I am comfortable using computers.

Strongly agree Agree Don't agree or disagree Disagree Strongly disagree

3. I am good at using computers.

Strongly agree Agree Don't agree or disagree Disagree Strongly disagree

4. I learn new computer skills quickly.

Strongly agree Agree Don't agree or disagree Disagree Strongly disagree

5. I use the Internet:

Every day 2-3 times a week Once a week 2-3 times a month Once a month
Less than once a month

6. I use the Internet to search for information:

Every day A few times a week Once a week A few times a month Once a month
Less than once a month

7. I can usually find what I am looking for on the Internet.

Strongly agree Agree Don't agree or disagree Disagree Strongly disagree

8. I have used a site like Del.icio.us or Flickr that uses tags to search for and organize information.

Yes No Not sure

Sample

Appendix C: Measuring Instrument

User Number: _____

Part 1

Please circle the answer that describes your feelings best.

1. The search system was easy to use.

Strongly agree Agree Don't agree or disagree Disagree Strongly disagree

2. It took me a long time to learn how to use the search system.

Strongly agree Agree Don't agree or disagree Disagree Strongly disagree

3. I often got frustrated while searching.

Strongly agree Agree Don't agree or disagree Disagree Strongly disagree

4. It was easy to figure out what to search for.

Strongly agree Agree Don't agree or disagree Disagree Strongly disagree

5. Most of the sources I found were useful for my paper.

Strongly agree Agree Don't agree or disagree Disagree Strongly disagree

6. I got a lot of results that had nothing to do with what I was looking for.

Strongly agree Agree Don't agree or disagree Disagree Strongly disagree

7. I was able to find a large number of useful sources.

Strongly agree Agree Don't agree or disagree Disagree Strongly disagree

8. It was easy to find information about my topic.

Strongly agree Agree Don't agree or disagree Disagree Strongly disagree

Part 2

Please take the time to fill out answers to these questions. Be as detailed and specific as you can. The information you give us will help us make the searching system better and easier to use for future students.

1. What was your research topic? _____

2. What were some of the problems you had while searching?

3. What (if anything) did you like about the searching system?

4. What (if anything) could have been better about the searching system?
