Folksonomies and Social Tagging: An Annotated Bibliography

Laura Baumann
INFO 522: Information Access and Resources
March 20, 2013
Introduction and Scope

The following bibliography discusses the use of folksonomies and social tagging in searching for information in today’s society. Many of these articles examine an array of studies that have been done about how effective folksonomies are in research, as well as touching on the positives and negatives that surround social tagging. Some of these studies compare folksonomies to the Library of Congress Subject Headings, while others try to find ways to predict how users might tag different sources using algorithms. Since folksonomies are a more recent development in research and locating information, the articles covered in this bibliography are quite current ranging from 2007-2012. Many of these articles were published in America, however several were published in Europe, and another in Canada.

Description

“Folksonomies are the vocabularies that result from the act of users’ application of subject terms, or tags, to particular items using their own vocabulary and understanding of, or relationship to, the object (DeZelar-Tiedman, 2011).” Traditionally, controlled vocabularies have been the basis of thesauri in library catalogs, however social tagging has become, in recent years, a more widely accepted form of cataloging and retrieving information. Controlled vocabularies follow strict guidelines and rules and are quite scientific, while social tagging is determined by users who create the tags (Wooline, Pickup, Udma, 2011). There is a large debate whether
this type of tagging is useful or a hindrance, which is why many studies and measures have been taken to evaluate this newer tool in library cataloging.

**Summary of Findings**

The main idea that kept arising from each of these articles and studies that were conducted was that folksonomies are a helpful tool in searching for relevant information, but are most effective when paired with controlled vocabularies such as the Library of Congress Subject Headings.

The biggest problems that arise from controlled vocabularies is that because they are indexed using a strict set of guidelines, the subject headings tend to be

**Bibliography**

**Entry 1:**


**Abstract:**

“Results of this exploratory study suggested engaging students in digital image tagging can have analytical and educational importance. The study was designed to gauge middle school students’ capacities to describe digital images from two digital libraries that they used in an information literacy activity. When describing the image attributes, students (N=81) freely chose single words and multiword phrases to describe the interpretations, feelings, and questions evoked by the images. These descriptors were used to derive conceptual categories for the seventeen digital images. Results demonstrated that students acknowledged the responsibility of indexers to choose index terms for objects in collections that enable identification, organization and retrieval. The study sheds light on the potential to improve age-appropriate access to images by means of offering a multi-tiered approach to image representation. It also introduces a transparent approach to teaching information literacy concepts through creative thinking about the meaning of resources and their relationship in a broader information cycle context.”

**Annotation:**
This study was unique and useful in discovering how the younger generation thinks critically. Middle school students were required to tag images and chose to index these objects. It was helpful to see the wide variation and similarities in the tags displayed by these seventh graders. You can see from this study that tagging trends hold true over multiple generations, and how social tagging of images is diverse due to varying interpretations of the same image.

Search Strategy:

I went directly to an advanced search on Hagerty Library’s website to find the article that Dr. Ercegovac mentioned that she had written on my topic. Several articles came up, but the third one appeared to be on my topic and was written in the year that was mentioned by Ercegovac so I selected it and was able to locate the full article through ebscohost.

Database:

Ebscohost

Method of Searching:

Author Search

Search String

Ercegovac, Zorana (au)

Scholarly/Refereed Status:

After locating this article through Ebscohost, it could also be found in Library Literature & Information Science Full Text database, which contains scholarly journals from 1982-present. Additionally, this journal was found to be peer-reviewed.

Entry 2:


Abstract:

“Academic libraries have moved toward providing social networking features, such as tagging, in their library catalogs. To explore whether user tags can enhance access to individual literary works, the author obtained a sample of
individual works of English and American literature from the twentieth and twenty-first centuries from a large academic library catalog and searched them in LibraryThing. The author compared match rates, the availability of subject headings and tags across various literary forms, and the terminology used in tags versus controlled-vocabulary headings on a subset of records. In addition, she evaluated the usefulness of available LibraryThing tags for the library catalog records that lacked subject headings. Options for utilizing the subject terms available in sources outside the local catalog also are discussed.”

Annotation:

Social tags and how they have entered libraries in the academic realm in recent years is detailed through this article. The author mentions LibraryThing and its usefulness of user tags. Tiedman explains how tagging has become so popular because users are able to collaborate with other users without realizing it and gives in depth examples of certain articles that have been tagged by users, which make the results of her study fully understandable in how they relate to user tags.

Search Strategy:

I selected Proquest, because I knew there are a wide variety of scholarly articles related to library science. I chose a keyword approach since this was one of my initial database searches beginning with just one of my keywords.

Database:

Proquest

Method of Searching:

Keyword

Search String:

“Social tagging” and library

Scholarly/Refereed Status:

This article is found in the Library and Resources and Technical Services Journal, a peer-reviewed journal, which is a scholarly journal of the Association for Library Collections Technical Services (ALCTS), a division of the American Library Association (ALA).

Entry 3:

Matthews, Brian; Jones, Catherine; Puzon, Bartlomiej; Moon, Jim; Tudhope, Douglas; et al (2010). An Evaluation of Enhancing Social Tagging with a
Abstract:

“Traditional subject indexing and classification are considered infeasible in many digital collections. This paper seeks to investigate ways of enhancing social tagging via knowledge organization systems, with a view to improving the quality of tags for increased information discovery and retrieval performance. Design/methodology/approach -- Enhanced tagging interfaces were developed for exemplar online repositories, and trials were undertaken with author and reader groups to evaluate the effectiveness of tagging augmented with control vocabulary for subject indexing of papers in online repositories. Findings -- The results showed that using a knowledge organisation system to augment tagging does appear to increase the effectiveness of non-specialist users (that is, without information science training) in subject indexing. Research limitations/implications -- While limited by the size and scope of the trials undertaken, these results do point to the usefulness of a mixed approach in supporting the subject indexing of online resources. Originality/value -- The value of this work is as a guide to future developments in the practical support for resource indexing in online repositories. Adapted from the source document.”

Annotation:

Ways in which user tags can be more effectively used by themselves through certain methods was discussed through this article. One particularly interesting method allows readers and authors' tags to be separated and used to create folksonomies that are more accurate for each article of information.

Search Strategy:

After reading the article, “Exploring User-Contributed Metadata's Potential to Enhance Access to Literary Works: Social Tagging in Academic Library Catalogs,” by Christine DeZelar-Tiedman, I looked through her citations and was able to access this journal article also located in Proquest.

Database:

Proquest

Method of Searching:

Citation

Search String:

Referenced in:

**Scholarly/Refereed Status:**

This article was found in the Aslib Proceedings, a scholarly peer reviewed journal of the Association for Information Management in the United Kingdom, which is an association for those who manage information and knowledge.

**Entry 4:**


**Abstract:**

“The purpose of this paper is to investigate the linking of a folksonomy (user vocabulary) and LCSH (controlled vocabulary) on the basis of word matching, for the potential use of LCSH in bringing order to folksonomies. Design/methodology/approach - A selected sample of a folksonomy from a popular collaborative tagging system, Delicious, was word-matched with LCSH. LCSH was transformed into a tree structure called an LCSH tree for the matching. A close examination was conducted on the characteristics of folksonomies, the overlap of folksonomies with LCSH, and the distribution of folksonomies over the LCSH tree. Findings - The experimental results showed that the total proportion of tags being matched with LC subject headings constituted approximately two-thirds of all tags involved, with an additional 10 percent of the remaining tags having potential matches. A number of barriers for the linking as well as two areas in need of improving the matching are identified and described. Three important tag distribution patterns over the LCSH tree were identified and supported: skewedness, multifacet, and Zipfian-pattern. Research limitations/implications - The results of the study can be adopted for the development of innovative methods of mapping between folksonomy and LCSH, which directly contributes to effective access and retrieval of tagged web resources and to the integration of multiple information repositories based on the two vocabularies. Practical implications - The linking of controlled vocabularies can be applicable to enhance information retrieval capability within collaborative tagging systems as well as across various tagging system information depositories and bibliographic databases. Originality/value - This is among frontier works that examines the potential of linking a folksonomy, extracted from a collaborative tagging system, to an authority-maintained subject heading system. It provides exploratory data to support further advanced mapping methods for linking the two vocabularies.”

**Annotation:**
The link between the LCSH and social tags is explored in this study. The LCSH tree is key in categorizing tags into folksonomies, showing how both professionally indexed subject headings and social tags can work together to find the most applicable information in a search.

**Search Strategy:**

I selected Proquest, because I knew there are a wide variety of scholarly articles related to library science. I chose a keyword approach since this was one of my initial database searches beginning with my other keyword.

**Database:**

Proquest

**Method of Searching:**

Keyword

**Search String:**

folksonomy

**Scholarly/Refereed Status:**

This article was found in the scholarly and peer reviewed *Journal of Documentation* after being located through the above steps in Proquest.

**Entry 5:**

Kipp, Margaret E.I. (2009). Information Organization Practices on the Web: Tagging and Social Organization of Information. The University of Western Ontario (Canada). *ProQuest Dissertations and Theses*, 391

**Abstract:**

“This thesis examines the phenomenon of social tagging from its early beginnings to its current level of prominence on a wide variety of websites in a series of linked studies examining the structures and patterns of tag term use to determine whether regular patterns appear that would support information organisation and retrieval.

The first study was designed to assess how collaborative tagging supports and enhances traditional ways of classifying and indexing documents. Statistical analysis of tagging patterns shows that many tagging practices are significantly similar or at least complementary to existing subject organisation systems. The
tags also indicated intriguing practices relating to personal information management which suggest an extra contextual dimension to information organisation which conventional systems do not generally facilitate.

The second study examined the context of online indexing from the viewpoint of the three different groups involved in information organisation: users, authors, and intermediaries. The results show that there are important differences in the context of keywords from the three groups.

The third study compared participant search behaviour using a social bookmarking tool and a more traditional journal database in order to determine if users found tags useful in their search process. The study showed that users did indeed make use of tags in their search process. Users also made use of controlled vocabularies in the journal database to locate useful search terms and links to related articles supplied by the database. Librarians did not previously think users would be willing to classify things and in general people thought user classifications would be too divergent or too personal to be useful; however, this study has shown that while user indexed systems have all of the flaws one might expect of such systems they also provide an intriguing exploration of user indexing and a strong support for indexing structures of all kinds in the support of information retrieval."

Annotation:

Kipp’s dissertation performed three very important studies. The study that was most helpful examined tagging from three objective perspectives to give an overall analysis of the usefulness of indexing or tagging. The different point of views helps to understand where each group is coming from and how they arrive at their varying tags.

Search Strategy:

I selected Proquest, because I knew there are a wide variety of scholarly articles related to library science. I chose a keyword approach since this was one of my initial database searches combining both of my keywords this time.

Database:

Proquest

Method of Searching:

Keyword

Search String:

“Social tagging” and folksonomy (keyword)
Scholarly/Refereed Status:

Through Proquest, and after refining my search strategy to peer reviewed information, I was able to locate this dissertation from a dissertation from the University of Western Ontario.

Entry 6:


Abstract:

"The purpose of this paper is to introduce the notion of social search as a new concept, drawing upon the patterns of web search behaviour. It aims to: define social search; present a taxonomy of social search; and propose a user-centred social search method. Design/methodology/approach: A mixed method approach was adopted to investigate and conceptualise the notion of social search. A review of literature on social search was carried out to identify the key trends and topics. A model of online collaboration was adopted to delineate the types and categories of social search. Four use case scenarios were developed to provide a more pragmatic approach to the understanding of social search. Findings: The developed taxonomy of social search reveals important similarities and differences between many social search systems. This analysis reveals a gap in social search approaches. A practical method was identified that allows users to directly leverage social search without special features built into search engines. Research limitations/implications: For feasibility reasons, Google was used as an example of a search system that can be used for carrying out social searches. Practical implications: The paper provides several practical implications for web searchers as well as web designers. In particular, some recommendations are provided for the design of search engines, digital libraries and browser add-ons. Social implications: The study demonstrates the value and power of "collective intelligence" in web search. It shows how general web searches can be enhanced through using socially enhanced web-based tools such as social bookmarking systems, social tagging services and social media sites. Originality/value: This is the first study that provides a granular analysis of the notion of social search and puts forward a taxonomy of social search. The use cases developed and reported are created based on real search topics, and show the value and validity of the approach taken."

Annotation:

This study in this article examines the taxonomy of social searching and the patterns that arise from it. It does not directly focus on social tagging, but it is
stated that social tagging plays a significant influence on social search behaviors, and discusses how social tagging has affected social searching.

**Search Strategy:**

I selected ERIC, because I knew there was a wide variety of educational documents and research that would support my topic. I chose a keyword search without using any quotations or other tools to hopefully find some general articles and reports on the topic.

**Database:**

ERIC

**Method of Searching:**

Keyword

**Search String:**

Social tagging (kw)

**Scholarly/Refereed Status:**

This article was found in the peer reviewed, refereed journal, Program: Electronic Library and Information Systems. It covers the management and use of information technology in libraries, archives, museums, and galleries. This article was located through the database ERIC.

**Entry 7:**


**Abstract:**

“This study explores the relationships between cultural and social capital and online social tagging behaviour in Delicious.com, a social bookmarking web site that offers social tagging functionalities. Based on Bourdieu’s conception of cultural and social capital, an online questionnaire was developed to measure Delicious users’ capital possession and its influences on social tagging behavioural tendencies. The study findings showed that the offline/online cultural capital and offline social capital affected information organization-oriented tagging; offline/online social capital affected social oriented-tagging; offline/online cultural capital and offline/online social capital both affected strategic tagging; offline/online social capital affected tagging imitation. Based on the findings, we
made inferences on the user roles and the power structure of a social tagging folksonomy community."

**Annotation:**

This was a very thorough study conducted to find the effects of offline/online social capital and online/offline cultural capital as well as possession of online/offline social/cultural capital. The results of the study were able to successfully make inferences about how users tag and create folksonomies within the online community.

**Search Strategy:**

I decided to try my keyword search in a different database specifically catered to articles about the library and information sciences and to see if I could find some new results with a similar search strategy. I also use Boolean logic in my keyword search so that I would find articles that may contain information on my subject in different endings.

**Database:**

Library Literature & Information Science Full Text (H.W. Wilson)

**Method of Searching:**

Keyword

**Search String:**

“Folksonomy” or “folksonomies”

**Scholarly/Refereed Status:**

This article comes from a peer-reviewed international journal of high recognition covering various issues and topics pertaining to library science. In addition, it was also located using the database Library Literature and Information Science, which holds a vast amount of information about the library sciences.

**Entry 8:**


**Abstract:**
“Social tagging, as a recent approach for creating metadata, has caught the attention of library and information science researchers. Many researchers recommend incorporating social tagging into the library environment and combining folksonomies with formal classification. However, some researchers are concerned with the quality issues of social annotation because of its uncontrolled nature. In this study, we compare social tags created by users from the LibraryThing website with the subject terms assigned by experts according to the Library of Congress Subject Headings (LCSH). The purpose of this study is to examine the difference and connections between social tags and expert-assigned subject terms and further explore the feasibility and obstacles of implementing social tagging in library systems. The results of our study show that it is possible to use social tags to improve the accessibility of library collections. However, the existence of non-subject-related tags may impede the application of social tagging in traditional library cataloguing systems.”

Annotation:

This unique article is applicable, because the study that was done, shows how social tagging and subject headings differ and overlap. The article also began by giving a very detailed description of how folksonomies came about and how they work as well as how subject headings are formed. It was helpful to be able to compare both professionally chosen subject headings to user chosen tags to see how the minds of professionals work in comparison to the reader’s.

Search Strategy:

Using Web of Science, a citation index, and keyword searching, I combined both key terms connected with Boolean logic to find articles that contained both of my terms.

Database:

Web of Science

Method of Searching:

Keyword

Search String:

Folksonomy and “social tagging”

Scholarly/Refereed Status:

This article comes from a peer-reviewed international journal of high recognition covering various issues and topics pertaining to library science.
Entry 9:


Abstract:

“Folksonomies emerge as the result of the free tagging activity of a large number of users over a variety of resources. They can be considered as valuable sources from which it is possible to obtain emerging vocabularies that can be leveraged in knowledge extraction tasks. However, when it comes to understanding the meaning of tags in folksonomies, several problems mainly related to the appearance of synonymous and ambiguous tags arise, specifically in the context of multilinguality. The authors aim to turn folksonomies into knowledge structures where tag meanings are identified, and relations between them are asserted. For such purpose, they use DBpedia as a general knowledge base from which they leverage its multilingual capabilities.”

Annotation:

This source was helpful, because it showed the difficulties that can arise from the use of social tags and folksonomies, but the author explains that there are knowledge bases such as DBpedia that are working to form folksonomies into a more meaningful and relational tool for finding information. This article did not just state the cautions of user tags, but went into great detail about how there are people working to make social tagging more effective and reliable through systems.

Search Strategy:

After difficulty in finding adequate information by searching both terms on Web of Science, I scaled my search back to one key term in hopes of finding better sources.

Database:

Web of Science

Method of Searching:

Keyword

Search String:
Folksonomy

**Scholarly/Refereed Status:**

This article was located through Web of Science, which is a citation index containing much scholarly work. Additionally, this article can be found in a journal that is edited by professors at universities around the globe.

**Entry 10:**


**Abstract:**

“The authors examine the differences between keyword searching and controlled subject headings in conducting research and the relative issues of each. They highlight the habits of students who frequently use Google to search and how this colors the students' ability to search in controlled environments like traditional databases. The benefits of taxonomies (the structure of LCSH) are examined as a means of more effective searching. Folksonomies and social tagging are also discussed. The authors highlight that LCSH is more in line with effective information literacy instruction and conclude that controlled subject heading searching is truly a more effective means of conducting research.”

**Annotation:**

The focus in this article was more on keyword searching and controlled subject headings in conducting research than folksonomies and social tagging. The author touches on how beneficial these two things can be when conducting research, but also cautions users about the inconsistencies that can often occur. This article is helpful in that it defines and covers the basic reasoning behind the use of folksonomies and social tags.

**Search Strategy:**

I used the database Library Literature and Information Science, because I knew the information I found would be pertinent to my subject, which is why therefore, I used a subject search.

**Database:**

Library Literature & Information Science Full Text (H.W. Wilson)

**Method of Searching:**

Subject
Search String:
Folksonomy(SU) or folksonomies(SU)

Scholarly/Refereed Status:
This article was located on the reliable database Library Literature and Information Science using a search selecting only peer reviewed journals.

Entry 11:

Abstract:
“This article will discuss the potential role of folksonomies and social tagging in the information control of scholarly articles. The article reviews claims that folksonomies may replace traditional indexing, criticisms of folksonomies and suggestions for their improvement. The primary conclusion is that, although folksonomies may not replace traditional thesaurus-based indexing, social tagging, as a means of both organizing scholarly articles and by drawing together groups of scholars interested in the same, and especially emergent, fields, can provide a useful method of information control by means of scholarly communication. The conclusions draw upon the 1968 book by Patrick Wilson Two Kinds of Power: An Essay in Bibliographic Control.”

Annotation:
Articulated in this article, is the role of folksonomies and social tagging in control of scholarly articles. The authors point out that folksonomies and social tagging cannot entirely replace other indexing, but it does help in scholarly communication. This claim is supported by the article, “Two Kinds of Power: An Essay in Bibliographic Control,” written by Patrick Wilson which the article focuses on.

Search Strategy:
I used the database Library Literature and Information Science, because I knew the information I found would be pertinent to my subject, which is why therefore, I used a subject search.

Database:
Library Literature & Information Science Full Text (H.W. Wilson)
Method of Searching:

Subject

Search String:

Folksonomy(SU) or folksonomies(SU)

Scholarly/Refereed Status:

By using a database that allows for searching only peer related scholarly journals, this article was located in the Canadian Journal of Information and Library Sciences.

Entry 12:


Abstract:

“As the web has grown into an integral part of daily life, social annotation has become a popular manner for web users to manage resources. This method of management has many potential applications, but it is limited in applicability by the cold-start problem, especially for new resources on the web. In this article, we study automatic tag prediction for web pages comprehensively and utilize the predicted tags to improve search performance. First, we explore the stabilizing phenomenon of tag usage in a social bookmarking system. Then, we propose a twostage tag prediction approach, which is efficient and is effective in making use of early annotations from users. In the first stage, content-based ranking, candidate tags are selected and ranked to generate an initial tag list. In the second stage, random-walk re-ranking, we adopt a random-walk model that utilizes tag co-occurrence information to re-rank the initial list. The experimental results show that our algorithm effectively proposes appropriate tags for target web pages. In addition, we present a framework to incorporate tag prediction in a general web search. The experimental results of the web search validate the hypothesis that the proposed framework significantly enhances the typical retrieval model.”

Annotation:

Analyzed in this article, are the ways in that users tag information, and showed how in this study, an algorithm was created to predetermine how a user might tag a certain article. You are able to see from the user perspective how they
formulate ideas in their head and assign certain tags to articles. It was neat to see how consistent and accurate the algorithm was.

Search Strategy:

Using the Library Literature and Information Science database, where you can find peer reviewed scholarly articles pertaining to the library science field, I performed a keyword search using the subject search.

Database:

Library Literature and Information Science Full Text (H.W. Wilson)

Method of Searching:

Keyword

Search String:

“Social tagging” or “social tags”

Scholarly/Refereed Status:

This article was located through a reliable database in the Journal of the American Society for Information Science and Technology that is a fully refereed and technical periodical and has been for over half a century.

Entry 13:


Abstract:

“This article examines the linguistic structure of folksonomy tags collected over a thirty-day period from the daily tag logs of Del.icio.us, Furl, and Technorati. The tags were evaluated against the National Information Standards Organization (NISO) guidelines for the construction of controlled vocabularies. The results indicate that the tags correspond closely to the NISO guidelines pertaining to types of concepts expressed, the predominance of single terms and nouns, and the use of recognized spelling. Problem areas pertain to the inconsistent use of count nouns and the incidence of ambiguous tags in the form of homographs, abbreviations, and acronyms. With the addition of guidelines to the construction of unambiguous tags and links to useful external reference sources, folksonomies could serve as a powerful, flexible tool for increasing the user-friendliness and interactivity of public library catalogs, and also may be useful for
encouraging other activities, such as informal online communities of readers and user-driven readers' advisory services.”

Annotation:

How folksonomies are formed and compared to tags from multiple sites and how well they align with the NISO guidelines is examined in this article. It was a very meticulous 30-day study, which in the end gave a helpful account of the problems associated with folksonomies as well as the usefulness they possess.

Search Strategy:

I decided to search through the references of an article that I had already located through Proquest to see if they had cited any sources that would also pertain to my subject.

Database:

Proquest

Method of Searching:

Citation

Search String:

Referenced in:

Scholarly/Refereed Status:

You can know that this is a reliable source, because this article was referenced in another article that has already been cited that is scholarly. It is also listed as peer reviewed in the Proquest database

Conclusion and Personal Statement

I knew next to nothing about what folksonomies were and what social tagging meant going into this project. They had been mentioned in another class and I was intrigued to know more. I learned that social tagging is a great way to
label information from a user perspective, because often times indexes are outdated or are too narrow in focus. Folksonomies are different than social tagging in that they are created through the tags that are given. They are viewed almost as an umbrella that covers certain tags since with social tagging, there are significantly more tags than when professionals choose the indexed terms.

I felt like I learned a lot about information structure not only from this assignment but also from what I learned about my topic. By learning about folksonomies, I also had to learn about how information was and is traditionally structured in catalogs. This helped me think about how I should word my search to achieve the best results, which I believe towards the end of my searching, I became well accustomed to.

I learned from searching that one term or selection could drastically change the results of a search. Just by changing the Boolean word “and” to “or” narrowed my search considerably. I also learned that there is so much information available that even performing the same exact search terms on a different database may achieve completely different results. I performed several searches that appeared almost identical in set up, but was able to locate a variation in results, which aided in my research.

This assignment was difficult for me. I had been struggling all quarter trying to become familiar with the various database and how they differ and become comfortable using Boolean logic and other techniques. I learned very quickly that waiting until halfway through the semester to start this project was a bad idea. To properly and adequately create and annotated bibliography, it takes time to carefully read each source and make sure that it is helpful. Even though the requirement is 12-20 sources, you end up reading more than that when you are wading through all that information. Annotated bibliographies are not that difficult to create as I thought it would be, as much as they are time consuming, however very helpful when conducting research.

I certify that:
• This assignment is entirely my own work.
• I have not quoted the words of any other person from a printed source or website without indicating what has been quoted and providing an appropriate citation.
• I have not submitted this assignment to satisfy the requirements of any other course.

Signature  ______Laura Baumann____________________
Date  _______3-20-2013_________________