Social Cataloging; social catalogers

Dr. Shawne Miksa
Social Tagging and its Effects on Catalogs
Cataloging and Metadata Roundtable
Texas Library Association Annual Conference
Houston, Texas
April 18, 2012
Questions being asked

- How can social or collaborative tagging be incorporated into traditional information systems such as library catalogs?
- What is the effect of social tags on the catalog?
- How does it affect the cataloger’s work?
  - How does it aid in subject cataloging and in particular subject analysis?
  - How does it affect authority control?
- How does it affect the catalog user?

How will the potential of social tagging best be harnessed? How can social tagging and vocabulary control interact? How does our concept and practice of authority control butt up against its complete opposite? How can we deliberately lose control over a time-honored process of authority control?

The introduction of the Internet and the World Wide Web to our professional world has leveled the field in such a way that the librarian is not the sole voice, but simply one among the many. How does this happen? If we place social tagging within the process of subject analysis and subject representation then might we simply equate social tagging to the brainstorming of an indexer or classificationists during the initial stages of the subject analysis process? As is currently practiced, the analytical process starts with examining a resource for keywords or phrases that represent the intellectual content. These are then translated into the language used in a controlled vocabulary. If this process can be aided by social tags, then how do we best take advantage of them? Alternatively, could we say that social tags are another species of controlled vocabulary in and of itself? Are the users doing our job for us and, if so, how well are they doing it?
At the same time we must also ask if the popularity of social tagging comes simply from the need or desire for simplicity of subject/content representation or ease of use/least effort, or perhaps even as a result of lack of understanding of controlled vocabularies? Is it born out of frustration of trying to understand and navigate a thesaurus, or can we assume it is simply a desire of users to gloss over the details in favor of rapid scanning of keywords as a quicker end to the angst of information need?

Is it born out of frustration of trying to understand and navigate a thesaurus, or can we assume it is simply a desire of users to gloss over the details in favor of rapid scanning of keywords as a quicker end to the angst of information need? Or, is it just a need to have an opinion? Is tagging a narcissistic act or an act of sharing knowledge? For example, subject analysis involves identifying underlying concepts within a resource in the hopes of producing a list of words that will serve in bringing together one particular information resource with all others of a similar subject matter, in addition to providing subject access for the user. How does this particular goal figure in the popularity of an individual, untrained user assigning their own terms to the resource (i.e., is this her goal?)
Fu, and co-authors give us a simple short definition: short textual labels

Voss (2007) states that “Tagging is referred to with several names.....the basic principle is that end users do subject indexing instead of experts only, and the assigned tags are being shown immediately on the Web.”

Tennis (2006) sees social tagging as type of indexing that is very personal.
I found that while social tagging as an activity started earlier, the bulk of research on tagging and catalogs, etc., within LIS really starts in earnest in 2006. There have been earlier studies, but they were broader and tended to focus on bookmarking or what was then simply called user-generated or use-created content added to ILS.
In many ways the study of social tags and tagging is similar to how the cataloging community reacted to ‘websites’ in the mid- to late 90s. The first instinct is to ask “What is it?” and then study the attributes, perhaps dissecting it---like you would a frog in biology class---in order to figure out how to define it, to compare it to the type, or species, of information resources with which we are already familiar.


Trant (2009) remarks that “This is a new areas of research, where theoretical perspectives and relevant research methods are only now being defined. This paper focuses on folksonomy itself. Three broad approaches are identified, focusing first, on the folksonomy itself (and the role of user tags in indexing and retrieval); secondly, on tagging (and the behaviour of users); and thirdly, on the nature of social tagging systems (as socio-technical frameworks).
Example studies

  - Weaknesses: “...potential for ambiguity, polysemy, synonymy, and basic level variation as well as the lack of consistent guidelines for choice and form...” (p. 23).

  - “This is a new areas of research, where theoretical perspectives and relevant research methods are only now being defined.
  - 3 broad approaches ...on folksonomy itself (and the role of user tags in indexing and retrieval)...tagging (and the behavior of users)...nature of social tagging systems (as socio-technical frameworks)...”

Trant (2009) remarks that “This is a new areas of research, where theoretical perspectives and relevant research methods are only now being defined. This paper focuses on folksonomy itself. Three broad approaches are identified, focusing first, on the folksonomy itself (and the role of user tags in indexing and retrieval); secondly, on tagging (and the behaviour of users); and thirdly, on the nature of social tagging systems (as socio-technical frameworks).
Tags and subject headings

- Rolla, P. (2009). User tags versus subject headings: Can user-supplied data improve subject access to library collections?
  - “A comparison of LibraryThing’s user tags and LCSH suggest that while user tags can enhance subject access to library collections, they cannot replace the valuable functions of controlled vocabulary like LCSH. (p. 182).

  - “…total proportion of tags being matched with LC subject headings constituted approximately 2/3 of all tags involved, with additional 10% of remaining tags having potential matches.” (abstract)

- “...many libraries are beginning to see tagging as a viable means of harnessing the wisdom of crowds (i.e., users) to shed light on popular topics and resources and involve users in collaborative, socially networked ways of organizing and retrieving resources.” (p. 57)
- “...tagging is a user-empowering way to organize and retrieve material, and many libraries have found it useful to incorporate this practice into their catalogs. The familiarity of tagging from other parts of the Internet has the potential to attract users who might otherwise be intimidated by more traditional catalog practices and structures.” (p. 58)
This is difficult for many catalogers to come around to, or agree with. It goes against our training in many ways.

We are trained to be objective when analyzing and assigning controlled terms to resources, which is exactly the opposite of how social tagging is used. The reader applies words and phrases that result out of their personal interaction and interpretation of a resource, and not necessarily with the broader audience in mind. The latter of which is exactly how most catalogers’ have been educated. We are trained to apply Haykin’s fundamental concept of ‘reader as the focus’ (1956), i.e., the cataloger’s own personal view is suspended in favor of reaching as broad an audience as possible. In essence, let the reader have her say; let the reader have a voice.
Steele points out many of the same weaknesses as Spiteri (2007), shown previously, there is a lack of hierarchy, no guarantee of coverage, synonymy, polysemy (more than one meaning), user’s intent.

An interesting piece of data: I asked a librarian at a public library that uses BiblioCommons how many tags have been added to their records---in the last 12 months around 3000 tags had been assigned, as compared to almost 100,000 ratings being done.
“Spagging”

- Arch (2007). Creating the academic library folksonomy: Put social tagging to work at your institution.
  - “There are, of course, a few risks and issues to consider when implementing social tagging in your library, especially if the site is open for all library patrons to update. One is the wonderfully named *spagging*, or spam tagging. Users with bad intentions can tag unsuitable sites for their own profit or simply to create havoc.” (p. 81)
There are dimensions to social tags that provide new food for thought when it comes to what we are representing.

Kipp & Campbell conducted a study of people searching a social bookmarking tool that specialized in academic articles. They found that while the participants used the tags in their search process, they also used controlled vocabularies to locate useful search terms and links to select resources by relevance.

“This study examined the relationship between user tags and the process of resource discovery from the perspective of a traditional library reference interview in which the system was used, not by an end user, but by an information intermediary who try to find information on another’s behalf.” (p. 252)

“The inclusion of subjective and social information from the taggers is very different from the traditional objectivity of indexing and was reported as an asset by a number of participants. The study suggests that while users value social and subjective factors when searching, they also find utility in objective factors such as subject headings. Most importantly, users are interested in the ability of systems to connect them with related articles whether via subject access or other means.” (abstract)
In terms of information behavior—they found that while participants had preferences for reducing an initial list of returns, or hits—adding terms, quick assessments, modify search based on results, scanning—but they were willing to change their search behavior slightly based on number of results. There was evidence of uncertainty, frustration, pausing for longer periods of time, hovering, scrolling up and down, confused by differences between controlled vocabularies and tags. They stated “It was fairly common for participants to use incorrect terminology to identify their use of terms when searching.” (p. 249).

Meaning that they don’t see clicking on a subject hyperlink the same as searching using a subject term.
This was an experiment supported in part by a grant from NSF, Office of Naval Research, and University of Illinois at Urbana Champagne.

A controlled experiment in which they directly manipulated information goals and the availability of social tags to study their effects of social tagging behavior in order to understand if content (i.e., semantics) of tags plays a critical role in tagging behavior. Previous research seemed to suggest that stabilization in tag choices are caused by two main factors: (a) the information goals of the users (i.e., what the user is looking for), and (b) the social influence of tags (i.e., how tags created by others influence future tag choices).

“Although previous models did implicitly assume the social influence of tags on other users as the major reason for the formation of emergent structures in the system, what is still lacking is direct evidence demonstrating the social influence of tags, and what is the nature of this social influence. It is still not clear, for example, whether there are other variables, such as differences in information goals, that moderate the social influence.” (p. 12:4)

It is still not clear, for example, whether there are other variables, such as differences in information goals, that moderate the social influence. To this end, we
designed an experiment to compare tagging behavior of two groups of users who can and cannot see tags created by others when using a social tagging system.
Ending with more questions

- What is the socialism aspect and do we, as information organizers, respect that aspect?
- For example, do people tag because they want to people to know who they are; does tagging give a person a way to say ‘this is who I am’?
- If a community of users all agree on a vocabulary then does it become a controlled vocabulary?
Social cataloging; social cataloger

**Working definition**
- The joint effort by users and catalogers to interweave individually- or socially- preferred access points, which can be both subject-based and task-based, with traditional controlled vocabularies in a library information system for the purpose of highly relevant resource discovery as well as user-empowerment. Both the user and the cataloger exercise their voice as to how information resources are related;
- An information professional/librarian who is skilled in both expert-based and user-created vocabularies, who understands the motivations of users who tag.

This is my working definition of social cataloging and social cataloger.
Final thoughts

- “...the importance of deciding aright where any given subject shall be entered in is inverse proportion to the difficulty of decision.” Cutter, C. (1904)

We need attempt to address some of these broader questions in the hopes of outlining a clearer process for the indexer or cataloger to follow when creating and providing intellectual access. Ultimately, I think it will perhaps convince catalogers to become more social catalogers then they have ever been in the past.

Cutter says it best here: The importance of making the right decision is inversely proportional to the difficulty of the decision.

One of the final effects of social tagging is the very obvious fact that cataloging work is not diminishing, but rather becoming more intense, and much more necessary to the library as a whole.