

# Folksonomies - Cooperative Classification and Communication Through Shared Metadata

Adam Mathes

Computer Mediated Communication - LIS590CMC  
Graduate School of Library and Information Science  
University of Illinois Urbana-Champaign

December 2004

## **Abstract**

This paper examines user-generated metadata as implemented and applied in two web services designed to share and organize digital media to better understand grassroots classification. Metadata - data about data - allows systems to collocate related information, and helps users find relevant information. The creation of metadata has generally been approached in two ways: professional creation and author creation. In libraries and other organizations, creating metadata, primarily in the form of catalog records, has traditionally been the domain of dedicated professionals working with complex, detailed rule sets and vocabularies. The primary problem with this approach is scalability and its impracticality for the vast amounts of content being produced and used, especially on the World Wide Web. The apparatus and tools built around professional cataloging systems are generally too complicated for anyone without specialized training and knowledge. A second approach is for metadata to be created by authors. The movement towards creator described documents was heralded by SGML, the WWW, and the Dublin Core Metadata Initiative. There are problems with this approach as well - often due to inadequate or inaccurate description, or outright deception. This paper examines a third approach: user-created metadata, where users of the documents and media create metadata for their own individual use that is also shared throughout a community.

## **1 The Creation of Metadata: Professionals, Content Creators, Users**

Metadata is often characterized as “data about data.” Metadata is information, often highly structured, about documents, books, articles, photographs, or other items that is designed to support specific functions. These functions are usually

to facilitate some organization and access of information. Administrative, structural, and descriptive metadata are three broad categories of metadata (Taylor, 2004). This paper focus primarily on descriptive metadata which identifies and functions to organize information based on its intellectual content.

Traditionally metadata is created by dedicated professionals. Catalogers create metadata, often in the form of Machine-Readable Cataloging (MARC) records for books and other intellectual creations, and this is the basis of most Online Public Access Catalogs (OPAC) in libraries and other institutions. This often requires serious education and training. The library and information science field has developed elaborate rules and schemes for cataloging, categorization and classification that include classification schemes such as the Dewey Decimal System and Library of Congress Classification Scheme, as well as large controlled vocabularies of terms for describing the subject of materials, such as the Library of Congress Subject Headings.

While professionally created metadata are often considered of high quality, it is costly in terms of time and effort to produce. This makes it very difficult to scale and keep up with the vast amounts of new content being produced, especially in new mediums like the World Wide Web. An alternative is author created metadata. Original creators of the intellectual material provide metadata along with their creations. The Dublin Core Metadata Initiative has been used with some success in this area (Greenberg et al, 2002). Author created metadata may help with the scalability problems in comparison to professional metadata, but both approaches share a basic problem: the intended and unintended eventual users of the information are disconnected from the process.

User created metadata is a third approach, and this paper focuses on grassroots community classification of digital assets. Other forms of user created metadata are often implicit. Citation analysis is a well established technique used to determine relationships between academic works and the impact of scholars. Similar analysis of the link structure in the World Wide Web is used by the PageRank algorithm, which became the theoretical basis for the Google search engine (Page, 1998). Recommendation systems, and those that employ collaborative filtering are another form of leveraging implicit user created metadata. (Lieberman, 2002).

One form of explicit user created metadata was popularized in the late 1990's with link-focused websites called weblogs (Blood 2000). These sites provide links combined with commentary, and are a form of mostly unstructured, but explicit, user created metadata. Customer reviews on web sites such as Amazon.com are an integral aspect of online commerce, and leverages consumer created metadata to create sites that are far more informative than comparable commercial sites.

## 2 Tagging Content in Del.icio.us and Flickr

Del.icio.us (<http://del.icio.us>, henceforth referred to as "Delicious") is a tool to organize web pages. A description online states it is:

“a social bookmarks manager. It allows you to easily add sites you like to

your personal collection of links, to categorize those sites with keywords, and to share your collection not only between your own browsers and machines, but also with others” (Schachter, 2004)

Delicious is not unique or pioneering in its role as bookmarks manager. What seems to be relatively new and different is the emphasis on user added keywords as a fundamental organizational construct. These keywords, which are referred to as “tags” on the site, allow users to describe and organize content with any vocabulary they choose.

To use the system, you must first join by registering an account. The system is free to join and use. Only a username, full name, and password are required. The user then adds a specialized bookmark to their web browser. When browsing a web page they would like to add to delicious, they select the bookmark, and are presented with a form that allows them to enter any tags they want to associate with the page, and then save it. These tags are optional; users can and do use the site without tagging their documents.

In addition to automatically generated chronological ordering of bookmarks saved to the system, the tags are used to collocate bookmarks within a user’s collection. Additionally, these tags are also used to collocate bookmarks across the entire system, so for example, looking at the page <http://del.icio.us/tag/linux> will show all bookmarks that are tagged with “linux” by any user.

Flickr (<http://www.flickr.com>), a photo management and sharing web application, has a similar system of free-form tagging for photos that was adopted and modeled after Delicious. It too requires users to create a user account, and is free to join. There is also the option to pay for an account with more features, like more storage space for photographs. Flickr offers a similar bookmark to add photographs to the system, but also has a number of other options to upload photographs to the system through web pages and software applications. Tags can be added at the time of upload, or later in the process when the photographs are displayed by the system.

A primary difference between Delicious and Flickr is that while the tags on Delicious are primarily from the users of web documents that were written by another party, Flickr is primarily used by individuals to manage their own digital images, and the majority of the tags are users tagging photos they created themselves. This is not absolute; the system does have the option of allowing users designated as friends or family to tag a users’ photos. Additionally, users can and do enter images others created into the system, often from web sites. This use of the system is much more like Delicious, but seems to be a small fraction of the use.

### 3 From Tags to Folksonomy

The most popular tags used on Delicious are listed on the right side of the front page. Related tags, as determined programmatically by the system, are listed on the right side of individual tag pages. The organic system of organization developing in Delicious and Flickr was called a “folksonomy” by Thomas Vander

Wal in a discussion on an information architecture mailing list (Smith, 2004). It is a combination of “folk” and “taxonomy.”

An important aspect of a folksonomy is that is comprised of terms in a flat namespace: that is, there is no hierarchy, and no directly specified parent-child or sibling relationships between these terms. There are, however, automatically generated “related” tags, which cluster tags based on common URLs. This is unlike formal taxonomies and classification schemes where there are multiple kind of explicit relationships between terms. These relationships include things like broader, narrower, as well as related terms. These folksonomies are simply the set of terms that a group of users tagged content with, they are not a predetermined set of classification terms or labels.

In Delicious, a cursory analysis of the tags reveals that the most popular tags are primarily subject descriptor keywords at various levels of specificity. Some of the most popular tags (as of November 14, 2004) according to the system were: “software, design, programming, music, politics, web, news, blog, css, linux, art, osx, java, mac, blogs, reference, fun, python, games, tech, photography, humor, tools, delicious, rss, firefox, toread, comics.” Many of these are technical subject tags reflecting the common interests of a tech-savvy user base, e.g. “rss, firefox, python, java, linux.” Some are best described as genre or form descriptors, like “comics, humor, fun, photography.” At least one, “toread,” is something qualitatively different: it is a tag apparently used for self-organization and reminder. Similarly, “wishlist” (<http://del.icio.us/tag/wishlist>) was apparently used by a number of users to highlight consumer items they were interested in.

The 150 most popular tags on Flickr are tabulated and listed on the site. As of November 19, 2004, this list included much of what one might expect as common subjects of photos: cat, friends, dog sky, sea, park, kids, garden, baby, building, flower, flowers signs, sculpture, city, vacation. Over 25% (41 out of 150) of the tags were proper place names like cities or countries. Colors were listed: yellow, green, blue, pink, orange, white, red. Years were also popular as tags, 2001, 2002, 2003 and 2004 were present amongst the most popular.

Some terms that have particular meaning in the photographic domain like portrait, macro, landscape, blackandwhite were included. The terms “cameraphone, moblog, fotolog” reflect the use of relatively new words, and the connection the site has to tech savvy early adopters of integrated camera phone technology and weblogs focused on photographs.

Two tags of particular interest are “cute,” and “me.” This will be elaborated upon later, but I think these two terms reflect the dual nature of these systems: the compulsion to share - what is the Internet if not a venue for sharing cute photographs? - and conversely the importance of individuality and ego for these systems to work.

Overall, although the term “classification” is often used in relation to these systems, and has been used in this paper, what is going on is more like “categorization.” Categorization is generally less rigorous and boundaries are less clear. It is based more on a synthesis of similarity than a systematic arrangement of materials (Jacob 2004). Most importantly, each document can have many terms associated with it. By contrast, classification schemes generally focus on pro-

viding a single classification to an item, and are very hierarchical and have clear relations. In a folksonomy the set of terms is a flat namespace: there are no clearly defined relations between the terms in the vocabulary.

### 3.1 Limitations

The problems inherent in an uncontrolled vocabulary lead to a number of limitations and weaknesses in folksonomies. Ambiguity of the tags can emerge as users apply the same tag in different ways. At the opposite end of the spectrum, the lack of synonym control can lead to different tags being used for the same concept, precluding collocation.

### 3.2 Ambiguity

As an uncontrolled vocabulary that is shared across an entire system, the terms in a folksonomy have inherent ambiguity as different users apply terms to documents in different ways. There are no explicit systematic guidelines and no scope notes. For example, items tagged with “filtering” on Delicious included the following:

- Last.FM - Your personal music network - Personalized online radio station
- InfoWorld: Collaborative knowledge gardening
- Wired 12.10: The Long Tail
- Oh My God It Burns! “ Practical Applications of the Philosopher’s stone. For drunks. Brita filter makes bad vodka into good vodka
- Introduction to Bayesian Filtering

These are all “filtering,” but in very different senses. Using water filters to purify vodka is a very different subject than Bayesian statistical analysis.

Acronyms present another area of potential ambiguity that are often dealt with effectively in controlled vocabularies. Examining the front page on November 14, 2004 revealed one user tagging sites with “ANT.” After examining the other sites the user tagged with ANT, it was apparent this was an acronym for “Actor Network Theory,” in the domain of sociology. However, when examining the ANT tag across all users (Delicious apparently is not case sensitive in tags) most of the bookmarks were about Apache Ant, a project building tool in the Java programming language. Two completely separate domains and ideas are mixed together in the same tag.

### 3.3 Spaces, Multiple Words

Both Delicious and Flickr seem designed primarily to deal with single words. Delicious does not allow spaces in tag names, although Flickr does. In some instances, multiple words are used together in a single tag, without spaces,

i.e., vertigovideostillsbbc' on Flickr. At times this can reflect users trying to put a hierarchy into a single tag, or simply reflects a category that has multiple terms, such as design/css' on Delicious. (<http://del.icio.us/tag/design/css.>) Both systems ignore letter case, which may collapse distinct ideas into a single tag, especially with acronyms.

### 3.4 Synonyms

There is no synonym control in the system. This leads to tags that seemingly have similar intended meanings, like “mac,” “macintosh,” and “apple” all being used to describe materials related to Apple Macintosh computers. Different word forms, plural and singular, are also often both present. In this particular situation with these Macintosh tags, the “related tags” sidebar of Delicious interlinks all three of these categories automatically. Plural vs. singular is often a problem, as seen in the popular tags on Flickr, both “flower” and “flowers” were listed.

These sorts of problems are the reasons why controlled vocabularies are used in many settings. Generally, any of the classic problems that controlled vocabularies help deal with will be present in these systems to varying degrees. However, it is likely that a controlled vocabulary would be impossible in the context of systems like Delicious and Flickr.

### 3.5 Strengths

Although a folksonomy is not a controlled vocabulary, and certainly does have limitations, there are important strengths that are important to understanding the appeal and utility of such systems.

### 3.6 Browsing vs. Finding

The first is serendipity. While the controlled vocabulary issues discussed above may hamper findability, browsing the system and its interlinked related tag sets is wonderful for finding things unexpectedly in a general area. In researching this paper, exploring the bookmarks tagged with “folksonomy” on Delicious, there were many recent resources from a wide variety of authors and sites that I likely would never have been exposed to.

There is a fundamental difference in the activities of browsing to find interesting content, as opposed to direct searching to find relevant documents in a query. It is similar to the difference between exploring a problem space to formulate questions, as opposed to actually looking for answers to specifically formulated questions. Information seeking behavior varies based on context. While one could evaluate a folksonomy in a system like Delicious or Flickr by using specific queries from users, and then evaluating which documents tagged with keywords they choose are relevant to the query, that would ignore the broader set of browsing activities that the system seems to be stronger in. Measuring the utility of that aspect would likely require qualitative research in the

form of interviews or ethnographic study of users, and is an area of further study. It would also require comparisons not to search based information retrieval systems, but to browsing activities using other categorization and classification schemes.

### 3.7 Desire Lines

Perhaps the most important strength of a folksonomy is that it directly reflects the vocabulary of users. In an information retrieval system, there are at least two, and possibly many more vocabularies present (Buckland, 1999). These could include that of the user of the system, the designer of the system, the author of the material, the creators of the classification scheme; translating between these vocabularies is often a difficult and defining issue in information systems. As discussed earlier, a folksonomy represents a fundamental shift in that it is derived not from professionals or content creators, but from the users of information and documents. In this way, it directly reflects their choices in diction, terminology, and precision.

Some classification schemes are disjoint from the vocabulary of the users. In “Metadata for the Masses,” Peter Merholz argues that a folksonomy can be quite useful in that it reveals the digital equivalent of “desire lines” (Merholz, 2004). Desire lines are the foot-worn paths that sometimes appear in a landscape over time. Merholz notes, “A smart landscape designer will let wanderers create paths through use, and then pave the emerging walkways, ensuring optimal utility. Ethnoclassification systems can similarly emerge.’ Once you have a preliminary system in place, you can use the most common tags to develop a controlled vocabulary that truly speaks the users’ language.”

Merholz recommends using a folksonomy as the start of professionally designed controlled vocabularies. While this may not be practical or desirable in many situations, the fundamental point is that the vocabulary of users may simply be too different from the other parties to adequately “pave the paths” in advance. Another important point may be that the terms users want to use move too quickly, or are qualitatively different than authors or systems designers.

Merholz’s own piece provides an excellent example. Merholz does not use the term “folksonomy.” He has written on his personal web site that the term is inaccurate due to its derivation from “taxonomy,” which he argues tend towards hierarchy and control. (Merholz, 2004) (See also Taylor, 2004, for discussions of problems and disputes with the term “taxonomy.”) Merholz prefers the term “ethnoclassification,” which is what he uses in his article, and there is no mention of “folksonomy” to be found. Ethnoclassification is also inaccurate, because as discussed, what is happening is quite unlike classification and far more like categorization.

Despite Merholz’s personal preference as author, his piece is tagged on Delicious with both “ethnoclassification” and “folksonomy,” as well as various other tags including “userexperience,” “tagging,” “taxonomy,” “metadata,” “social-software,” and “facets.” The tags reflect not the author’s vocabulary, or any

particular classification or categorization system’s vocabulary, but the language and vocabulary that individual users choose to describe the article with.

Although the Delicious tags on Merholz’s article are only one example, a folksonomy, with its uncontrolled nature and organic growth, has the capability to adapt very quickly to user vocabulary changes and needs. There is no significant cost for a user or for the system to add new terms to the folksonomy. The problem is that while the disparate user vocabularies and terms enable some very interesting browsing and finding, the sheer multiplicity of terms and vocabularies may overwhelm the content with noisy metadata that is not useful or relevant to a user.

## 4 Why Folksonomies Work

It is difficult to define a metric by which one could argue folksonomies are a success or failure, but the degree that it does seem to be effective in these systems as a way of organizing information, and that a large group of people are using these systems, I posit, is due to a few important factors. The overall costs for users of the system in terms of time and effort are far lower than systems that rely on complex hierarchal classification and categorization schemes. In addition to this structural difference, the context of the use in these systems is not just one of personal organization, but of communication and sharing. The near instant feedback in these systems leads to a communicative nature of tag use.

### 4.1 Barriers to Entry, Cognitive Costs

The conceptual shift from professional, designed, clearly defined categorization and classification schemes to an ad-hoc set of keywords enables users not just professionals without any training or previous knowledge to participate in the system immediately. Additionally, participating is far easier in terms of time, effort and cognitive costs.

Stewart Butterfield, one of the creators of Flickr, argues that the difference in complexity between folksonomies and classification schemes is important:

“Aside: I think the lack of hierarchy, synonym control and semantic precision are precisely why it works. Free typing loose associations is just a lot easier than making a decision about the degree of match to a pre-defined category (especially hierarchical ones). It’s like 90% of the value of a proper’ taxonomy but 10 times simpler.” (Butterfield, 2004)

Many professionals would likely argue that Butterfield’s assessment of 90% and “10 times simpler” is vastly overstated, his fundamental point holds true: non-trivial and important metadata are captured through these folksonomies. The comparisons are almost irrelevant as it would be impossible to get the users of these systems to use a complex, hierarchical, controlled vocabulary. The barriers are simply too high. A system that tried to capture that full value

would cost too much in user time, effort, and cognitive cost, and thus have little value in practice.

## 4.2 Feedback and Asymmetric Communication

Jon Udell (2004) argues that the idea of abandoning taxonomy in favor of lists of keywords is not new, and that the fundamental difference in these systems is feedback.

“Of course, that idea’s been around for decades, so what’s special about Flickr and del.icio.us? Sometimes a difference in degree becomes a difference in kind. The degree to which these systems bind the assignment of tags to their use - in a tight feedback loop - is that kind of difference.

Feedback is immediate. As soon as you assign a tag to an item, you see the cluster of items carrying the same tag. If that’s not what you expected, you’re given incentive to change the tag or add another. If your items aren’t confidential and online-only access is sufficient, this can be a great way to manage personal information. But the real power emerges when you expand the scope to include all items, from all users, that match your tag. Again, that view might not be what you expected. In that case, you can adapt to the group norm, keep your tag in a bid to influence the group norm, or both.” (Udell, 2004)

This tight feedback loop leads to a form of asymmetrical communication between users through metadata. The users of a system are negotiating the meaning of the terms in the folksonomy, whether purposefully or not, through their individual choices of tags to describe documents for themselves.

There are two models to describe what is happening here: one that focuses on individual incentives, and one that focuses on community aspects.

## 4.3 Individual and Community Aspects

Both Delicious and Flickr are used by individuals to organize materials with their own vocabulary of terms. Individuals have an incentive to tag their materials with terms that will help them organize their collections in a way that they can find these items later. The organizational scheme that emerges for each individual reflects their individual information needs. The popularity of the “me” tag on Flickr perhaps best reflects this aspect of a folksonomy, as well as the “toread” tag on Delicious. Both can really only be understood in the context of an individual user.

Conversely, both Delicious and Flickr are services designed to share materials. The individual organizational behavior takes place in a public virtual space on these web sites. Therefore the behavior of the users can also be thought of as being influenced and related to their relationship to the other individuals using the service, and specific groups of users who they share tag use with. It is perhaps harder to justify this model simply from examination of the tags used, but there is definitely evidence of communication and perhaps even community formation through metadata, which will be discussed later.

A folksonomy lowers the barriers to cooperation. Groups of users do not have to agree on a hierarchy of tags or detailed taxonomy, they only need to agree, in a general sense, on the “meaning” of a tag enough to label similar material with terms for there to be cooperation and shared value. Although this may require a change in vocabulary for some users, it is never forced, and as Udell discussed, the tight feedback loop provides incentives for this cooperation.

Finally, there is the compulsion to share in general that underlies these systems. The very act of user self-selecting what to tag is important: this is not just material that users want to find themselves later, but also material they are sharing with others. Both systems have an explicit kind of social networking component built-in: Flickr allows you to specify other users as contacts, friends, or family and see views of just their material; Delicious allows you to “subscribe” to other users lists.

These two models, community and individual motivations, are not mutually exclusive, and it is likely both are necessary to explain a folksonomy in the context of these services. An area of further qualitative analysis could help to determine how much each of these theories applies to actual user behavior.

#### 4.4 Unanticipated Uses

While the folksonomies that developed at Flickr and Delicious have a definite focus on subject categorization, there are tags being used in some unexpected, interesting ways that reflect communication and ad-hoc group formation facilitated through metadata.

Flickr’s “**sometaithurts**” - for “so meta it hurts” is a collection of images regarding Flickr, and people using Flickr. (<http://www.flickr.com/photos/tags/sometaithurts/>) The earliest image is of someone discussing social software, and then subsequent users posting screenshots of that picture within Flickr, and other similarly self-referential images. The referential and meta nature of the images continues as users took pictures of images on Flickr, etc. Although this is a playful example, it is a use of tags as communicative tool. Only by tagging their photograph with “sometaithurts” could a user of the system join the photographic conversation. Conversely, the only way to follow the conversation was through the systems automated collocating of like tagged items.

A user on Flickr, Andrew Lowosky, began posting pictures of doorbells in Florence, along with a brief piece of fiction about the doorbell in the description of the photograph. He dubbed this combination of photograph and short story “**flicktion**,” and tagged it as such. (Lowosky, 2004.) Some other users have been tagging photographs with “flicktion” and writing short fiction to accompany it, although as of November 19, 2004, there were only three other users. Although small, there is a quick formation of new terms to describe what is going on, and others adopting that term and the activity it describes.

Examining all photos in Flickr tagged with “**iraq**” includes photographs Iraq, US troops in Iraq, as well as photographs of war protests. Although this may not be a community, what we are seeing is a group of people helping to define a term with their photographs and metadata.

## 5 Areas For Further Research

### 5.1 Quantitative Tag Analysis

Examining the quantitative aspects of folksonomies is an area that could yield some interesting data on the makeup and use of the terms used to describe items. One area to examine is the distribution of tag use: I hypothesize that it follows a power law scenario. That is, the most used tags are more likely to be used by other users since they are more likely to be seen, and thus there will be a few tags that are used by a substantial number of users, then an order of magnitude more tags that are used by fewer users, and another order of magnitude more used by only a handful of users. Examining this sort of distribution of tag use could give a better indication of whether a folksonomy converges on terms and foster consensus, or if as the user based grows the vocabulary grows at a more even rate, and the distribution of terms flattens, perhaps indicating less agreement.

### 5.2 Qualitative User Analysis

Examining user behavior through ethnographic observation or interview to understand user motivations and cognitive processes in tagging items would clarify what factors directly influence the formation of a folksonomy, and how individual incentives and group communication motivations influence use of the system. Although it seems that some users are intending to facilitate communication through tag use, especially in the unintended uses, interviews could make this point explicit. Interviews could also elucidate the conscious intentions of users in “normal” use of the system, which is much harder to observe simply from the documents and tags themselves. Other behavior that would be helpful to observe is the frequency with which users modify or change their tags, or future tagging behavior based on the implicit feedback from the system in the form of what other documents are tagged with a term.

### 5.3 Applicability to other systems

Delicious and Flickr are large web services designed to organize and share digital works. The applicability of user generated free-form tagging as an organizational construct in other contexts bears further investigation. One interesting area would be to systems where there is already an existing social network and examine how this related to the system. A corporate intranet, or a system in an academic setting used by a department’s faculty and students.

The use of a folksonomy to supplement existing classification schemes and provide additional access to materials by encouraging and leveraging explicit user metadata contributions is a possible area for research and further development in information retrieval systems. If information retrieval systems begin to incorporate user-centered information management tools, the organizational

schemes developed by the users have the possibility to be of great interest to other users and improve the systems.

## 6 Conclusion

A folksonomy represents simultaneously some of the best and worst in the organization of information. Its uncontrolled nature is fundamentally chaotic, suffers from problems of imprecision and ambiguity that well developed controlled vocabularies and name authorities effectively ameliorate. Conversely, systems employing free-form tagging that are encouraging users to organize information in their own ways are supremely responsive to user needs and vocabularies, and involve the users of information actively in the organizational system. Overall, transforming the creation of explicit metadata for resources from an isolated, professional activity into a shared, communicative activity by users is an important development that should be explored and considered for future systems development.

## 7 References

1. Blood, Rebecca. "Weblogs: A History and Perspective", Rebecca's Pocket. September 7, 2000. [[http://www.rebeccablood.net/essays/weblog\\_history.html](http://www.rebeccablood.net/essays/weblog_history.html)]
2. Butterfield, Stewart. "Sylloge." August 4, 2004. [<http://www.sylloge.com/personal/2004/08/folksonomy-social-classification-great.html>]
3. Buckland, Michael. "Vocabulary as a Central Concept in Library and Information Science." Digital Libraries: Interdisciplinary Concepts, Challenges, and Opportunities. Proceedings of the Third International Conference on Conceptions of Library and Information Science (CoLIS3, Dubrovnik, Croatia, 23-26 May 1999. Ed. by T. Arpanac et al. Zagreb: Lokve, pp 3-12. [<http://www.sims.berkeley.edu/buckland/colisvoc.htm>]
4. Greenberg, Jane, Maria Cristina Pattuelli, Bijan Parsia and W. Davenport Robertson. "Author-generated Dublin Core Metadata for Web Resources: A Baseline Study in an Organization." Journal of Digital Information, Volume 2 Issue 2 Article No. 78, November 2001. [<http://jodi.ecs.soton.ac.uk/Articles/v02/i02/Greenberg>]
5. Jacob, Elin K. "Classification and categorization: a difference that makes a difference." Library Trends. Winter, 2004. [[http://www.findarticles.com/p/articles/mi\\_m1387/is\\_3\\_52/ai](http://www.findarticles.com/p/articles/mi_m1387/is_3_52/ai)]
6. Lieberman, Henry. "Interfaces That Give and Take Advice." From Human-Computer Interaction in the New Millenium. John M. Carroll, Ed. Boston: Addison Wesley, 2002.
7. Losowsky, Andrew. "The doorbells of florence project." The Prandial Post. September 21, 2004. [[http://www.prandial.com/archives/2004\\_09.html#008740](http://www.prandial.com/archives/2004_09.html#008740)].

8. Merholz, Peter. "Ethnoclassification and vernacular vocabularies." August 30, 2004. [<http://www.peterme.com/archives/000387.html>]
9. Merholz, Peter. "Metadata for the Masses." October 19, 2004. [<http://www.adaptivepath.com/publication>]
10. Page, L.; Brin, S.; Motwani, R.; and Winograd, T. 1998. The PageRank citation ranking: Bringing order to the web. [<http://wwwdb.stanford.edu/backrub/pageranksub.ps>]
11. Schachter, Joshua. Del.icio.us About Page. [<http://del.icio.us/doc/about>]. 2004.
12. Smith, Gene. "Atomiq: Folksonomy: social classification." Aug 3, 2004 [[http://atomiq.org/archives/2004/08/folksonomy\\_social\\_classification.html](http://atomiq.org/archives/2004/08/folksonomy_social_classification.html)]
13. Taylor, Arlene G. The Organization of Information. Westport: Libraries Unlimited 2004.
14. Udell, Jon. "Collaborative knowledge gardening." InfoWorld. August 20, 2004. [[http://www.infoworld.com/article/04/08/20/34OPstrategic\\_1.html](http://www.infoworld.com/article/04/08/20/34OPstrategic_1.html)]