



PROJECT MUSE®

---

Folksonomies, Social Tagging and Scholarly Articles / Les folksonomies, l'étiquetage social et les articles scientifiques

David Woolwine, Margaret Ferguson, Eric Joly, David Pickup, Cristian Mihai Udma

Canadian Journal of Information and Library Science, Volume 35, Number 1, March/mars 2011, pp. 77-92 (Article)

Published by University of Toronto Press

DOI: <https://doi.org/10.1353/ils.2011.0002>



➔ *For additional information about this article*

<https://muse.jhu.edu/article/419606>

# Folksonomies, Social Tagging and Scholarly Articles

# Les folksonomies, l'étiquetage social et les articles scientifiques

---

David Woolwine  
Joan and Donald E. Axinn Library  
123 Hofstra University  
Hempstead, NY 11549-1230  
David.Woolwine@hofstra.edu

Margaret Ferguson, Eric Joly, David Pickup, and Cristian Mihai Udma  
School of Information Studies  
McGill University  
Montreal, QC

**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*.

**Keywords:** indexing, folksonomies, social tagging, information control, scholarly communication

**Résumé :** Cet article explore le rôle potentiel des folksonomies et de l'étiquetage social comme outils de gestion de l'information des articles scientifiques. L'article passe en revue les assertions selon lesquelles les folksonomies pourraient remplacer l'indexation traditionnelle, les critiques des folksonomies et les suggestions pour l'amélioration des folksonomies. La conclusion principale est que, même si les folksonomies ne vont pas remplacer l'indexation traditionnelle basée sur le thésaurus, l'étiquetage social, comme moyen d'organisation des articles scientifiques et d'assemblage des groupes de chercheurs intéressés par les mêmes et surtout les nouveaux terrains de recherche, peuvent représenter une méthode utile de gestion de l'information au moyen de la communication savante. Les conclusions s'inspirent du livre rédigé par Patrick Wilson en 1968 portant le titre *Two kinds of power : an essay in bibliographic control*.

**Mots-clés :** indexation, folksonomies, étiquetage social, contrôle de l'information, communication scientifique

## Introduction

This article will discuss the potential role of folksonomies and social tagging in the information control of scholarly articles. Specifically it will argue that, although folksonomies and social tagging cannot at present entirely replace other forms of information control (e.g., traditional indexing and algorithmic searching), social tagging may play an important role in pointing to emerging work in academic disciplines and can aid in scholarly communication. The focus of the article will be a discussion of Patrick Wilson's *Two Kinds of Power: An Essay in Bibliographic Control*. It is argued that Wilson has theoretically laid out the basis for ideal forms of information control and that he points to the need for supplements to impersonal forms of information control. This is especially the case for emerging areas of research and where connections between concepts have been made by only a small number of scholars. The primary conclusion is that, although social tagging may not replace traditional thesaurus-based indexing and algorithmic searching, it has a role to play in information control by facilitating scholarly communication and by drawing together groups of scholars interested in emergent fields.

The first three sections will summarize the literature on current use of folksonomies and social tagging. The summary of literature provides a context in which to introduce the main argument. In the literature review it is argued, first, that criticisms of social tagging from the perspective of traditional indexers are valid only if social tagging is seen as a type of indexing; second, that revolutionary claims that folksonomies have the potential to replace traditional indexing are not strongly supported and miss a potentially more important role for social tagging; and third, that step-wise improvements and more recent research in the field are promising but that the best way to conceive of social tagging systems are as social networks or social communication tools. This is where Patrick Wilson provides theoretical insight into the role for social tagging. The article will then take up Wilson and the conclusions that can be drawn from his work.

## Literature review: Traditional criticisms of folksonomies and social tagging

From the perspective of traditional indexing, several general and philosophical criticisms of social tagging as a means of information control

can be raised. One is that users' motivation for tagging cannot be known and can be infinitely diverse. Morrison (2007) argues that reasons for adding a tagging functionality to a site may be to improve search results, to add a navigational facet, to classify with minimal cost, or to get users to interact. Some of users' motivations to tag are to find things later, to get exposure and traffic, to voice their opinion, to take advantage of the functionality, to play a game, or to earn points. None of these motivations are remotely similar to those of traditional indexers.

Another criticism is that the use of tags to navigate is not equivalent to information retrieval. To what extent is navigation similar to information retrieval? Very often in an online environment such navigation means browsing without a precise direction or clear intention. O'Connor (1996, 65) states that "*browsing* is the activity or set of activities used by scholars to get around the difficulties of representing documents in advance, of use by searchers with no clearly stated goal." Indexing does not facilitate browsing as a rule. Furthermore, the purposes behind searching when motivated to find something precisely defined by the user and browsing may not be served by the same means of information control. When one searches by means of an index, one at least believes that one knows what to expect. When one is searching through socially tagging navigational systems, one seldom knows exactly what to expect as results. One only expects to get something. The precise "see" and "see also" cross-references are by no means equivalent to the cloud of tags suggested in these navigational systems. From a traditional indexing perspective, providing the user with "something" is not information retrieval. A systematic use of a social tagging system might mean first identifying all imaginable tagging motivations that might assist one in the search, then thinking about possible terms that might have been used, then performing a search, hoping to retrieve something useful and relevant. In other words, in the navigational environment provided by social tagging there is no way to obtain authoritative retrieval, which is defined as presenting something according to unbiased, systematic, and universal rules. In such an environment one can obtain only some orientation, or guidance, or in extreme cases only insights.

Finally, traditional indexing is focused on first determining the "aboutness" of a text and then affixing agreed-upon terms that capture or represent that aboutness. It is by these means that authoritative retrieval is believed to occur. This view, however, rests, at least in part, on the assumption that texts have stable meanings that experts can be trained to decode

(Rafferty and Hilderly 2006). Such a view holds that the correspondence between index terms and information documents is direct. But in social tagging the correspondence is established by “consensus” (Johncocks 2008). From a traditional indexing standpoint, this type of consensus does not mean objective, or at least established, truth.

This conservative defence of traditional indexing, which relies upon some version of a correspondence theory of truth and assertions of certainty of results, is unwarranted. The issue is not whether the index or indexer can establish a permanent or near-permanent correspondence between an indexing term and the aboutness of an article but rather how stable academic discourse is and to what extent an index employs and appropriately changes terms that are useable by those working within academic discourses. From this viewpoint, both indexing and social tagging are social activities, and the meanings of indexing terms and tags arise out of those activities. The real question becomes which set of social activities produces best access in service to the user: the activities associated with some versions of traditional thesaurus-based indexing, or those associated with social tagging/folksonomies? It should be kept in mind that users may have multiple and changing needs in their information-seeking behaviour.

Furthermore, the aboutness of an item indexed is always a matter of interpretation and there is never a final settled and perfect statement of what a document is “about.” It is important to point out that indexing terms themselves do not have stable meanings. The meanings of indexing terms change as at least two contexts change: the context of the index itself (a shift in the meaning of one term can lead internally to shifts in meanings of other terms) and the context of the academic discourse to which indexing terms are meant to provide access. As for “certainty” of results, what is certainty? “Certainty” that one knows what one is going to get when using a traditional index is largely a feeling of confidence and can be based on several factors. One is simply a belief in the authority of the index, often itself based on reputation or on the historical performance of a particular index. It also depends on what information or type of information the user is seeking. All of this can change over time. It can change because of changes in the performance of indexes, and it remains an empirical question whether social tagging, or algorithmic searching, or some combination of indexing and such other means, can provide or be made to provide confidence in results.

More centrally to the claims of this article, such criticisms are meaningful only if social tagging is viewed as a type of replacement for traditional indexing, if it is meant to either provide less expensive access to information with the same degree of satisfaction or even if it is meant as a supplement to traditional indexing in a sense of performing “another standard search.” What if social tagging were viewed as another form of information control altogether, as a supplement but of another type? What if dialogue, communities of practice, short-term consensus (not a type of “universal consensus”), various motivations, and some sort of browsing might be valued in some information-seeking contexts? The claim of this article is that such browsing is valued in some contexts and especially where no consensus has emerged. The weaknesses of social tagging and folksonomies from the perspective of traditional indexing then become a strength.

### **Literature review: Revolutionary claims**

Another way of looking at folksonomies is to see them as a form of *emergent indexing*. One could call it emergent indexing because aggregated tagging resembles networks and what has been called “the wisdom of crowds” (Surowiecki 2004) or *emergence* (Johnson 2001). As Mathes (2004) postulated, and as the work of Guy and Tonkin (2006) supports, tag distribution is scale-free in the manner of most human networks, and becomes increasingly so the more taggers join the system. Mathes himself prefers *ethnoclassification*, which has problems of its own as a term, despite its history of use in field anthropology (Rosenfeld 2005). Furthermore, as Golder and Huberman (2006) discovered, it takes only about 100 users tagging an item for the pattern to emerge. In other words, operating in a scale-free network environment, users’ binary actions (or simple rules: use an existing tag or add a new tag) create another network with a similar distribution. All this leads those who see folksonomy as a new form of information control to claim that tag agglomerations have the potential to lead to a consensus surprisingly quickly on the aboutness of any item. Since aboutness is a basic concept and problem of traditional indexing, the claim is then made that folksonomies could lead to a breakthrough, to a new way of doing indexing without the time-consuming, human-capital intensive ways of the past.

Furthermore, David Weinberger (2007) argues that in the digital world of information it is actually counterproductive to attempt to organize

items by rules and limitations that pertain in the material world. A good example is a book in a library, where, despite multiple subject headings, the physical object itself must be given a call number and put in a single space on a shelf. In the digital world, an infinite number of items can occupy the same space, while one item can be in infinite places. Weinberger would argue that this constraint is what leads to the prevalence of hierarchical taxonomical and thesaurus structures in the material world. But when dealing with digital items, using a structure designed to parallel physical constraints seems inappropriate. Shirky (2005) puts it simply, stating that “there is no shelf.” Additionally, the network patterns produced by aggregated tagging activity (where the nodes are the tags themselves and the links are co-occurrences) mirror the structure of digital networks (Cattuto et al. 2007). Golder and Huberman (2006) have postulated that consensus tags correspond to agreed-upon basic levels. However they deal with patterns emerging in the deli.cio.us site. This site is popular and non-specialized and has built-in procedures for collaborative tagging that push users toward consensus. Therefore, their work, and claims about easily emergent consensus of social tags, remains suggestive, especially in regard to scholarly articles.

Other work has empirically explored the use of social tags in the information control of scholarly works. This has largely involved comparisons of author-supplied keywords, social tags, and controlled vocabulary from thesaurus-based indexes. Kipp (2005) examined three types of information-control techniques (author-supplied keywords, social tags, and controlled vocabulary) on CiteULike, asking if terms and patterns of usage are similar or different for articles in two journals in information science. Kipp (2007) later carried out a similar but larger study using two journals in biomedicine. The major finding of this last study was that, in the comparison of tagging and indexing in the biomedical journals, taggers were more interested in methodology and “user groups” than were indexers. This was not the case in the journals in information science. Kipp also argues that tagging patterns may suggest the scholars think in terms of “associative trials,” but this conclusion is not strongly supported by the studies. Bruce (2008) compares indexing terms for articles in the Education Resources Information Center (ERIC) and social tags in CiteULike. His goal is to see the extent to which tags can exhibit patterns and provide an “emerging vocabulary.” Such studies—undertaken with the intent of determining the extent to which social tagging of scholarly articles is like more standard forms of information control (indexing, author-supplied keywords)—are also preliminary in

their conclusions. However, Kipp's studies point to the possibility that social tagging may be something other than a replacement for, or standard supplement to, the other methods of information control and therefore offers some support to the claims of this article.

Finally, Weinberger (2007) is incorrect in analogizing article indexing and (book) subject classification. The call number assigned to a book does, in effect, give a "primary subject" and slots it into a specific location on a shelf. But other subject headings also place the book, in a metaphorical sense, in other places. Historically this was in a card catalogue (along with author and title locations), but today in electronic catalogues its record can be viewed along with those of other items with identical—and often if there is a list display—similar subject headings. More importantly, article indexing never placed an item on a shelf. It always slotted items into multiple places (i.e., indexing terms often in various hierarchies) or gave, in another turn of phrase, multiple points of access. Indexing might be said to have always been halfway between the material universe and a virtual one, even when indexes themselves existed as material things, as printed books. The issue is not whether we can represent multiple locations in cyberspace. The real issue is what mix of social practices and tools serve the needs of users.

### **Literature review: Step-wise improvements**

Many of the suggestions for step-wise improvements of social tagging and folksonomies have involved tag displays. Hassan-Montero and Herrero-Solana (2006) have proposed an improved method for tag layout where tags are grouped by similarity based on clustering techniques and co-occurrence analysis. Their proposal for tag-cloud layout is based on the assumption that clustering techniques (bundles) can improve the experience of browsing tag clouds. Applying data clustering (grouping similar or linked tags) instead of listing them alphabetically would be a first step in structuring these data that appear at times to be chaotic, sloppy, and redundant.

Beyond display issues, a general suggestion has been made that automated techniques might be used to reduce the more obvious problems. For instance, Wu, Zhang, and Yu (2006) have also suggested methods for stripping out ambiguous terms. MacGregor and McCulloch (2006) have addressed the issues of ambiguity, polysemy, and synonymy. Misspellings of tags and mis-tagging might be addressed by having users



rate individual tags. Ontology-building software in general may remedy the lack of thesaurus-like hierarchical structures, although there is still a need for human intervention and labour. For instance, Heymann and Garcia-Molina (2006) propose an algorithm that would automatically convert a large corpus of tags into a navigable hierarchical taxonomy. Out of the noisy tag data generated by thousands of users would emerge taxonomies. The method remains a work in progress.

“Mash-ups” or floating a search engine over a system of social tags has also been suggested. Mash-ups of Flickr and del.icio.us tags with search sites already exist. Sinha (2006) lists three broad categories. The first is faceted tag browsing, as used by RawSugar, Siderean (fac.etiou.us), and mfeedia.com—the most successful of the types of mash-up. However, this still needs a huge amount of human intervention to work properly and thus the cost-benefit is not entirely clear. The second is clustering and recommenders (Flickr and del.ici.ous), which can occasionally promote intuitive leaps but are not conducive to power-use or serious focused research. The third is pivot browsing as exemplified by, among others, dogear.com and airtightinteractive.com’s Flickr tag browser. These are aesthetically pleasing to some users and promote resource discovery but do not support focused search. Finally, the work of Good, Kawas, and Wilkinson (2007) may prove promising in the application of social tagging to scholarly articles. They have created a mash-up (E.D. or entity describer) that connects the Connotea social tagging system, an index, and a Resource Description Framework (RDF) database for storing social semantic annotations and by which social taggers can intentionally connect their own tags with concepts in controlled vocabularies. This has, however, been carried out only in a life science context.

### **Two Kinds of Power and a way forward with social tagging**

In a remarkable book, *Two Kinds of Power: An Essay on Bibliographical Control*, Patrick Wilson (1968) lays out the requirements for ideal bibliographic instruments that allow for information control. In many ways Wilson anticipated the web-based mechanisms that are now taken for granted, including algorithm-based searching. But we will argue that he also anticipated a role for social tagging in information control.

In chapter 7, “Indexing, Coupling, Hunting,” Wilson makes several important distinctions. The first is between internal and external criteria.

Readers today may not want to entirely accept his position that texts have such boundaries. Today it is generally accepted that all texts are in some sense part of larger discourses, but the distinction, at least as far as it goes, is a useful one. Wilson writes, “Internal criteria are those whose application requires looking at nothing but the writing being judged; external criteria are those whose application requires looking beyond the writing itself” (98). He goes on to refer to the fact that indexing is not done, or is rarely done, for one person or even for only a few, but rather always for a group. The imagined or ideal interests—and today we would add shared meanings—of a group has always been something that an indexer keeps in mind in both constructing a thesaurus and applying indexing terms to an item at hand. Wilson also makes a further distinction between “impersonal external criteria” (what we might call a basic knowledge of the discipline, largely consensus knowledge and conservative in the sense of not opening up new issues or questions) and “personal” or “idiosyncratic” criteria (what we would call a private interest, perhaps not a private meaning but one rare enough to not be shared by many others).

Wilson believes that the user of an index always has unique concerns. Scholars as a rule are not only interested in uncovering what has been done but ideally have research interests directed toward the creation of something new. Therefore it is likely that an index created on the basis of internal and “impersonal external criteria” would, nonetheless, require a scholarly user to “engage in exploration.”

If users understand the hierarchical arrangements of the thesaurus, they can explore existing terms with relative ease. Visual displays of thesauri may make that type of exploration even easier in today’s online indexes. Abstracts are also useful, and Wilson even mentions the usefulness of “samples of the writing.” (In today’s databases not just samples but the full text of the article may often be looked over quickly by the user.) Abstracts, samples, and full-text displays allow users to begin to create their own classifications that are not based on the hierarchical arrangements in the thesaurus. But Wilson argues that the membership of such categories is based largely on guesses. Looking for documents to fit into those categories (well or badly constructed) is called by Wilson “hunting,” and selecting them he calls “picking.” Wilson is adamant that an index cannot anticipate the many interests that users will have and can never be constructed (using established thesauri, which are based on shared meanings and a general knowledge of the discipline) in a way that eliminates the need for such hunting and picking by scholars.

Wilson also argues that, beyond the standard “places” represented in a traditional index by terms, there exists the usual “auxiliary apparatus” that traditionally has been “see” and “see also” references. Wilson points out that this second apparatus has two purposes. It primarily guides the user through the complex hierarchy of a traditional thesaurus. He believes it also helps in hunting, providing “suggestions . . . to one who, not having found what he wanted, or all of what he wanted, at one position, is uncertain about where to go next. The customary apparatus consists of a system of what we call *couplings* of positions” (1968, 106). These would be “links” in today’s parlance. Such couplings are further divided between “analytic links,” which are species/genus or part/whole relations and therefore are purely logical and “synthetic links,” which depend upon knowledge of the “world.” (That steamships are ships is an analytic statement, that Pierpont Morgan is a banker or that diamonds are used in industry is a synthetic one.) One thing a good indexer does is assist the user in this type of navigating by supplying synthetic connections of which the user may be unaware. Wilson, however, goes on to discuss a third “sort of coupling,” which he calls “overlap” (1968, 107). This coupling or link is the degree to which content is shared and can be, in his view, very useful to users. It is probably useful to quote Wilson at length here. He writes,

The third sort of coupling might be called the “overlap” sort of coupling. To establish the first two sorts, one needs read no books at all, or no particular books; such couplings can be established simply on the basis of reflection on the meanings of the terms used in the rules of application, or on the basis of “common knowledge.” But one might establish a coupling of positions on the basis of examination of the writings assigned those positions, and of a discovery of a degree of overlapping content. Writings assigned position N1 might characteristically contain the same sort of talk as those assigned position N2, perhaps because of some “association” of topics that has little or no basis in factual or analytic connections of the topics discussed. Books assigned a position described in terms of, say, the history of Sanskrit literature might characteristically overlap in content with those assigned a position described in terms of the history of Indian medicine, where works “on” the history of Greek literature might characteristically *not* overlap with those “on” the history of Greek medicine, and this for, as we would say, “accidental” reasons of the history of scholarship. (1968, 107)

Today, algorithms provide a powerful auxiliary apparatus that gives users a good indication of this type of overlapping content in articles. Wilson, in the 1960s, has largely pointed to what we now have in algorithmic search engines. His footnote at this point in the chapter indicates that he is aware of merely formal statistical means of relating articles.

In regard to synthetic couplers, or “factual” relationships between terms, Wilson believed there was no way at the time he was writing to indicate the degree of closeness of relationship between terms. (For example, is horsemanship more or less closely related to stirrups than to blacksmiths?) Furthermore, given the nature of scholarship as a creative activity, anything might, in principle, be related to anything else. He writes,

But let us suppose that we could devise a method that would allow us to identify, for each subject or topic represented a repertory of positions, the five or ten or hundred most closely related positions, in order of their closeness of relationship. Suppose the related positions coupled in such a way that we could automatically pick out the position most closely related to any given position, then proceed to the next most closely related, and so on. What good would this do us? It must be that there is what we would call a “General Rule of Hunting” which goes something like this: Discussions of a thing X are more likely to be found in the context of discussions of a thing Y, the more closely Y is related to X. (1968, 110)

However, Wilson admits that the likelihood of a meaningful relationship between overlapping terms (even in closely related items) is not the same as certainty and is probably only good as a tool for someone knowing little of a field. A real scholar would find, or want, the other types of “overlap.” Wilson writes, “An auxiliary apparatus would be preferable that discovered and noted actual overlaps, that consisted of ‘overlap’ couplers rather than of analytic or factual ones” (1968, 111). These are those overlaps that are due to the “accidental” nature of scholarship but are of central importance to the scholar who is especially interested in areas where emerging connections are being made. Wilson believed that annotated bibliographies, made by experts in the field, were the most likely apparatus to point to such areas of scholarship in his time but also argued that a perfect apparatus of this sort cannot be created and that “hunting” will always be necessary for the researcher.

Is there a way to create such an apparatus in our day and one that does not rely upon annotated bibliographies? Wilson’s work may point the way and this is where social tagging comes in. In the chapter 8, “Consultants and AIDS,” he describes a social network of experts in a field or subfield that could evaluate an area of scholarship. Given the ever-changing nature of scholarship, such groups might be small but nonetheless in touch with one another. They would “circulate information” that feeds their interests and, even more importantly, monitor closely related fields to see if information or concepts could be imported (1968, 121). Wilson further contrasts such groups with established bibliographic

instruments such as thesaurus-based indexes. He writes of such established instruments, "But these public memories are rigid in organization and inflexible in description; they do not respond to questions, or only in limited and rigidly defined ways, and cannot reevaluate their own contents" (124). Arguing for the necessity of human flexibility he writes, "A human being, on the other hand, can reconsider his stock of knowledge and opinion in terms of a completely novel concept, or a completely novel requirement" (124). How can one produce a bibliographic instrument that allows for the discovery of novel concepts (some of which arise out of the relationship of previously unrelated concepts, and are, therefore, new or unexpected overlaps), the transmission of such concepts to others who may be interested, and at least some evaluation by experts? Furthermore, if such novel concepts are the result of new and rare overlaps, they would appear initially to be "bizarre" and of interest to only a few. Only the dedicated and uniquely motivated hunter would find such couplings and would sense the usefulness of such findings. They would be found most likely by a type of directed serendipitous browsing (Foster and Ford 2003). One can assume they would be found only in a relatively small number of articles and that they would most likely be distantly related and that therefore algorithmic search displays might not directly or immediately reveal them. Although such connections might be picked up by indexers, this is not guaranteed and, given the tendency of online indexes to display articles chronologically and to suggest only frequently co-occurring subjects, such connections might not be displayed or hinted at. If the concept arising from the overlap were a truly emerging one, no indexing term would exist, and furthermore indexers, who often "think in indexing terms," may not even see novelty. To add to Wilson's analysis of indexing, therefore, not only is the thesaurus conservative and "rigid," but the practices of indexers may be also, and the present online displays work against the revelation of novel connections.

Here we can return to one of Wilson's original distinctions, between "impersonal external criteria" and the idiosyncratic external criteria. Social tags are the best means for creators of scholarly content to signal and point to their idiosyncratic connections and interests. Furthermore, completely new interests and concepts are more likely to be assigned a new and unique tag. A hunter who is looking for such connections and concepts, or luckily coming upon them, would be greatly aided by such tags. So social tagging becomes something other than a weaker form of coupling provided by algorithms and something other than the

shared consensus meanings of established research represented usually by traditional indexing. It becomes a way of pointing to what are initially apparently bizarre connections and to areas of research where such connections may occur and to the emerging concepts themselves. Here information control intersects with scholarly communication, for one researcher seeing that another is making seemingly bizarre but interesting connections can, within such a network or community, begin to search on those tag combinations, or follow that scholar's work, and contact that individual and begin communication and collaborative tagging. The group of bibliographic consultants that Wilson calls for can then arise within such a social tagging network, for such scholars could build a community and create a new area of scholarly discourse. Indexers might also regularly examine tags for such instances and for the emergence of not only new "indexing terms" (which might represent new concepts) but also possible new subfields and new areas of research and note that concepts previously unrelated have started to co-occur.

Databases whose function is to allow access to scholarly articles would need to incorporate searching and display capacities that highlight unusual connections between items and not only, as some have argued, primarily power-tags (Peters and Stock 2010). And if the argument above is correct, controlled vocabulary would be less likely to point to truly emergent areas of research, and social tagging would be the best way to point to such areas. Determining whether this is, in fact, the case is one possible area of further research. Furthermore, use of social tags for information control and scholarly communication might solve the problem raised by Weinberg (1988): indexes do not well serve more advanced scholars because traditional indexes have trouble pointing to the "aspect" of articles. This too could be an area of future research. Kipp's work gives some indication that scholars in at least some fields are more likely to tag on aspect than are indexers. Finally, the view presented here from a theoretical perspective meshes well with work that explores social tagging as a product of communities of practice and as socially situated (Pfeiffer and Tonkin 2010; Tonkin et al. 2008). Whether members of particular disciplines or sub-disciplines are more likely to use social tagging to reach out to and communicate with other scholars working in emergent areas are other questions for future research.

Social tagging and collaboration based on social tagging among scholars has begun. Some examples are Springer's CiteULike, Firefox's Zotero,

Blackboard's Scholar, and Elsevier's 2collab. In a study by Elsevier on the use of social networking tools, it was reported that some researchers were using such tools (which included social networking practices other than social tagging of scholarly articles) and that they expected greater use in the future. "Find interesting information" was the primary reason listed for using social networking tools, and "research collaboration" was among the top five reasons given. Improvements desired among present users were for social networking to be able to find "more relevant information," "efficiently find specialists in my area of interest," and "view research output from a particular individual" (Elsevier 2008). The practices of collaborative sites are suggestive only, but the general view expressed here, that social tagging has a role in helping to locate unique, "bizarre connections," which is then used by active researchers/scholars in a field or subfield to create new concepts or products, may have begun to find some support in practice.

## References

- Bruce, R. 2008. "Descriptor and Folksonomy Concurrence in Education Related Scholarly Research." *Webology* 5 (3).  
<http://www.webology.ir/2008/v5n3/a59.html>.
- Cattuto, C., C. Schmitz, A. Baldassarri, V.D.P. Servedio, V. Loreto, A. Hotho, M. Grahl, and G. Stumme. 2007. "Network Properties of Folksonomies." *Artificial Intelligence Communications* 20 (4): 245–62.
- Elsevier. 2008. "STM Social Media Survey: Key Conclusions & Results." [http://www.2collab.com/static/files/2collab\\_social\\_media\\_survey.pdf](http://www.2collab.com/static/files/2collab_social_media_survey.pdf).
- Foster, A.E., and N.J. Ford. 2003. "Serendipity and Information Seeking: An Empirical Study." *Journal of Documentation* 59: 321–40.
- Good, B.M., E.A. Kavas, and M.D. Wilkinson. 2007. "Bridging the Gap between Social Tagging and Semantic Annotation: E.D. the Entity Describer." *Nature Precedings*. <http://precedings.nature.com/documents/945/version/2>.
- Golder, S.A., and B.A. Huberman. 2006. "Usage Patterns of Collaborative Tagging Systems." *Journal of Information Science* 32: 198–208.
- Guy, M., and E. Tonkin. 2006. "Folksonomies: Tidying Up Tags?" *D-Lib Magazine* 12 (1). <http://www.dlib.org/dlib/january06/guy/01guy.html>.
- Hassan-Montero, Y., and V. Herrero-Solana. 2006. "Improving Tag-Clouds as Visual Information Retrieval Interfaces." Paper presented at the International Conference on Multidisciplinary Information Sciences and Technologies, InSciT2006, October 25–8, 2006, Mérida, Spain.  
[http://nosolousabilidad.com/hassan/improving\\_tagclouds.pdf](http://nosolousabilidad.com/hassan/improving_tagclouds.pdf).

- Heymann, P., and H. Garcia-Molina. 2006. "Collaborative Creation of Communal Hierarchical Taxonomies in Social Tagging Systems." *InfoLab Technical Report 2006 10*: 1–5.
- Johncocks, B. 2008. "Web 2.0 and Users' Expectation of Indexes." *Indexer* 26: 18–24.
- Johnson, S. 2001. *Emergence: The Connected Lives of Ants, Brains, Cities, and Software*. New York: Scribner.
- Kipp, M.E.I. 2005. "Complementary or Discrete Contexts in Online Indexing: A Comparison of User, Creator, and Intermediary Keywords." *Canadian Journal of Information and Library Science* 29: 419–36.
- . 2007. "Tagging Practices on Research Oriented Social Bookmarking Sites." *Proceedings of the 35th Conference of the Canadian Association for Information Science*. [http://www.cais-acsi.ca/proceedings/2007/kipp\\_2007.pdf](http://www.cais-acsi.ca/proceedings/2007/kipp_2007.pdf).
- Macgregor, G., and E. McCulloch. 2006. "Collaborative Tagging as a Knowledge Organization and Resource Discovery Tool." *Library Review* 55: 291–300.
- Mathes, A. 2004. "Folksonomies: Cooperative Classification and Communication through Shared Metadata." Adam Mathes. <http://www.adammathes.com/academic/computer-mediated-communication/folksonomies.html>.
- Morrison, P.J. 2007. "Why Are They Tagging, and Why Do We Want Them To?" *Bulletin of the American Society for Information Science and Technology* 34: 12–15.
- O'Connor, B. 1996. *Explorations in Indexing and Abstracting: Pointing, Virtue, and Power*. Englewood, CO: Libraries Unlimited.
- Peters, I., and W.G. Stock. 2010. "'Power Tags' in Information Retrieval." *Library HiTech* 28. [http://www.phil-fak.uni-duesseldorf.de/infowiss/admin/public\\_dateien/files/56/1235723134power\\_tags.pdf](http://www.phil-fak.uni-duesseldorf.de/infowiss/admin/public_dateien/files/56/1235723134power_tags.pdf).
- Pfeiffer, H.D., and E.L. Tonkin. 2010. "eTagging in Context: Information Management across Community Networks." in *Handbook of Research on Social Interaction Technologies and Collaboration Software: Concepts and Trends*. ed. T. Dumova and R. Fiordo, 158–69. Hershey, PA: Information Science Reference.
- Rafferty, P., and R. Hilderly. 2007. "Flickr and Democratic Indexing: Dialogic Approaches to Indexing." *Aslib Proceedings: New Information Perspectives* 59: 397–410.
- Rosenfeld, L. 2005. "Folksonomies? How about Metadata Ecologies?" [LouisRosenfeld.com](http://louisrosenfeld.com). [http://louisrosenfeld.com/home/bloug\\_archive/000330.html](http://louisrosenfeld.com/home/bloug_archive/000330.html).
- Shirky, C. 2005. "Ontology Is Overrated: Categories, Links and Tags." Clay Shirky's Writings about the Internet. [http://www.shirky.com/writings/ontology\\_overrated.html](http://www.shirky.com/writings/ontology_overrated.html).
- Sinha, R. 2006. "Findability with Tags: Facets, Clusters and Pivot-Browsing." *Rashmi's Blog*. <http://rashmisinha.com/2006/07/27/findability-with-tags-facets-clusters-and-pivot-browsing/>.
- Surowiecki, J. 2004. *The Wisdom of Crowds: Why the Many Are Smarter Than the Few and How Collective Wisdom Shapes Business, Economies, Societies, and Nations*. New York: Doubleday.



Tonkin, E., E.M. Corrado, H.L. Moulaison, M.E.I. Kipp, A. Resmini, H.D. Pfeiffer, and Q. Zhang. 2008. "Collaborative and Social Tagging Networks. *Ariadne* 54 <http://www.ariadne.ac.uk/issue54/tonkin-et-al/>.

Weinberg, B.H. 1988. "Why Indexing Fails the Researcher." *Indexer* 16: 3–6.

Weinberger, D. 2007. *Everything Is Miscellaneous: The Power of the New Digital Disorder*. New York: Times Books.

Wilson, P. 1968. *Two Kinds of Power: An Essay on Bibliographic Control*. Berkeley: University of California Press.

Wu, X., L. Zhang, and Y. Yu. 2006. "Exploring Social Annotations for the Semantic Web." *Proceedings of the 15th International Conference on World Wide Web*, 417–26. New York: Association for Computing Machinery.