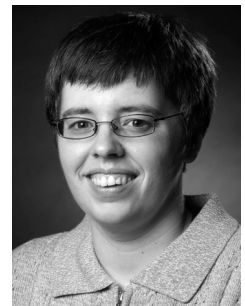# Tagging of Biomedical Articles on CiteULike: A Comparison of User, Author and Professional Indexing

Margaret E. I. Kipp

School of Information Studies, University of Wisconsin-Milwaukee, Bolton Hall Rm: 510, 3210 N Maryland Ave, Milwaukee, WI, USA 53211 <kipp@uwm.edu>

Margaret E. I. Kipp is an Assistant Professor and member of the Information Organization Research Group, School of Information Studies, University of Wisconsin-Milwaukee. She has a background in computer science and worked as a programmer/analyst. She has a PhD in Library and Information Science from the University of Western Ontario. Her research interests include social tagging, information organisation on the web, classification systems, information retrieval, collaborative web technologies and the creation and visualisation of structures in information organisation systems.

ABSTRACT: This paper examines the context of online indexing from the viewpoint of three different groups: users, authors, and professional indexers. User tags, author keywords, and descriptors were collected from academic journal articles, which were both indexed in PubMed and tagged on CiteULike, and analysed. Descriptive statistics, informetric measures, and thesaural term comparison shows that there are important differences in the use of keywords among the three groups in addition to similarities, which can be used to enhance support for search and browse. While tags and author keywords were found that matched descriptors exactly, other terms which did not match but provided important expansion to the indexing lexicon were found. These additional terms could be used to enhance support for searching and browsing in article databases as well as to provide invaluable data for entry vocabulary and emergent terminology for regular updates to indexing systems. Additionally, the study suggests that tags support organisation by association to task, projects, and subject while making important connections to traditional systems which classify into subject categories.

1.0 Introduction

The development of information organisation schemes is often related to significant increases in the size of document collections. The invention of writing and the subsequent recording of information created the first sets of documents that needed to be stored for later retrieval. While early writing and information access was restricted to the small group of educated citizens, mass education and mass production have created an increasing amount of information with a resultant interest in locating and using that information. As Eisenstein notes in *The Printing Revolution in Early Modern Europe*, the development of the movable type printing press in the 1500s caused an upsurge in the amount of printed material and provided a pressing need to increase the capacity of organisational systems for documents (Eisenstein 1983).

One such shift in the pace and volume of information production is occurring now, as academics and researchers increasingly turn to the web to locate articles, often in preprint archives. The increasing existence of open access archives, open access journals, and web archives of conference proceedings has increased the availability of research materials prior to and after publication and increased pressure on traditional indexing and organisation systems for organisation and retrieval. The substantial increase in access to information afforded by the Internet has only

246

M. E. I. Kipp. Tagging of Biomedical Articles on CiteULike: A Comparison of User, Author and Professional Indexing

Knowl. Org. 38(2011)No.3

strengthened the importance of being able to distinguish at once between similar documents and locate relevant documents.

The rise of collaborative tagging systems suggests an alternative method for creating indexing systems. In fact, such social bookmarking sites are sometimes touted as a potential solution to the problems of scale inherent in the application of any controlled vocabulary to a large document set and may have the potential to aid in providing the benefits of a controlled vocabulary, which controls for terminological differences, while still allowing the use of natural language vocabulary (Hammond et al 2005; Morville 2005; Shirky 2005). To discover if tags can truly provide a useful replacement or enhancement for controlled vocabularies, it is important to examine whether or not they provide a similar contextual dimension to existing indexing systems.

2.0 Social Bookmarking Tools

Social bookmarking is the act of sharing bookmarks by associating a URL with a username and a set of useful labels or tags. Since tags and social bookmarks are public, there is also the potential for sharing amongst groups of users using the same tags or bookmarking the same URLs. Social bookmarking sites have become increasingly popular since their inception in 2003 with delicious.com reporting 5.3 million users in 2008, the most recent statistic provided (http://blog.delicious.com/blog/2008/11/delicious-is-5.html).

CiteULike (http://citeulike.org/) is a social bookmarking service specialised for use by academics who wish to bookmark academic articles for later retrieval. CiteULike was created by Richard Cameron in November 2004 (http://www.citeulike.org/faq/faq.adp). Similar to delicious.com, CiteULike allows users to assign tags to the articles in their library. Users are encouraged to add as many or as few tags as they feel are necessary to help them relocate the article. Since many items bookmarked on CiteULike are journal articles, it is also possible to collect author keywords and descriptors assigned to these same articles. Thus, a comparison can be made between user tags, author keywords, and professional indexer descriptors attached to a single article.

3.0 Related Studies

Traditional indexing methods have tended to rely on trained indexers to organise and describe information.

While indexing of documents has a long history in library science, it has not been without controversy. Index terms and controlled vocabularies attempt to improve recall and precision by eliminating the ambiguity inherent in natural languages (and were also the only source of search terms before the advent of full text databases), but often require a large entry vocabulary to allow access to the controlled vocabulary. Users often express admiration of the controlled vocabulary systems, but are frustrated trying to match their own vocabulary to that of the thesaurus (Campbell and Fast 2004). Additionally, there is the issue of inter indexer consistency. Many studies have shown that inter indexer consistency is consistently quite low despite training and a shared context. These results hold true across various forms of media and in various fields (Markey 1984; Chan 1989).

Mathes (2004) notes that one important reason for continuing to seek new methods of generating index terms is the issue of scalability. While traditional professional indexing could be used to increase recall and precision, it is slow and expensive to apply. Automatic indexing is faster, but suffers from many of the same problems as free text search using natural language queries. Tagging is a potential bridge between these two different methods of indexing, allowing user terminology and decreasing the cost of indexing by accepting indexer input from the actual users of the documents.

Early studies of tagging (Hammond et al. 2005; Kipp and Campbell 2006) show that tags can be quite different from indexing terms, but some differences between user categories and indexer categories may be simply a matter of the use of standard techniques in indexing, such as the use of nouns instead of verbs or the elimination of plurals (Cleveland and Cleveland 1983, 101-102). Others may relate to the depth of the indexing itself, such as document level indexing versus exhaustive indexing. While users might find indexing of individual chapters of a book very useful, this would be extremely expensive. Additionally, since indexing is designed to be useful to the largest number of users possible, it is difficult to provide task specific indexing. Subject and topic indexing exposing the essential context or aboutness of the item has been the goal, allowing users to locate the item and add their own specific task related terms.

Kwasnik (1991) noted the importance of factors that are not subject related in the organisation of personal documents. This includes research documents and project related material. Terms such as @toread and cool are used in tagging (Kipp 2007a), but are

considered to be too short term, too user centric, or too subjective to be included in traditional subject heading or indexing systems. These short-term, user specific tags suggest important differences between tagging and professional indexing.

Mathes (2004) notes that there are three common groups involved in the assignment of keywords to documents. These three different groups—authors, professional indexers, and users—have distinct needs and purposes behind their indexing efforts and thus may be expected to use differing terminologies. Professional indexers, often librarians, assign controlled vocabulary subject headings to an article using their own domain knowledge and training. Journal authors may also be expected to assign keywords to their articles. In addition to these two traditional groups, a third group of indexers has arisen on social bookmarking sites. Users bookmarking articles on CiteULike are encouraged to tag articles for the purpose of organisation or retrieval.

Both title and author keywords have received relatively little attention in the literature. A few studies have examined author keywords (Schultz, Schultz, and Orr 1965; Kipp 2005; Gil-Leiva and Alonso-Arroyo 2007; Kipp 2007b; Strader 2009; Kipp 2011), while a few additional studies have examined title keywords (Bloomfield 1966; Voorbij 1989; Frost 1989; Ansari 2005, Jeong 2009).

Schultz, Schultz, and Orr (1965) examined author keywords and title keywords for 285 biomedical articles submitted for publication and found that author terms were more likely than title terms to match the controlled vocabulary terms. Kipp (2005) examined tags, author keywords, and descriptors and found that while many tags matched descriptors or author keywords exactly, there were also substantial differences. Gil Leiva and Alonso-Arroyo (2007) compared author keywords to descriptors using 640 abstracts from a variety of journal databases and found that 46% of author keywords matched descriptors when normalised. Strader (2009) compared author keywords and *LCSH* headings assigned to electronic theses. She found that approximately 37% of author keywords matched *LCSH* terms exactly and concluded that author keywords would make useful additional access points.

Bloomfield (1966) studied simulated machine indexing using title keywords compared to subject headings assigned to journal articles and found that 20% of title keywords and terms from the abstract matched subject headings. Voorbij (1998) compared title keywords of monographs in the humanities and social sciences and descriptors at the National Library of the Netherlands using thesaural relationships for comparison and found that, while only 10% of titles corresponded poorly with the content, descriptors were still an asset and resulted in the retrieval of more relevant results. Frost (1989) compared title keywords to Library of Congress Subject Headings and found that the degree of match (from 2-23% exact match between title and subject heading) was strongly dependent on the field of study (science and technical subjects had a higher degree of match). Ansari (2005, 414) examined the degree of exact and partial match between title keywords and descriptors of medical theses in Farsi. She found that the degree of match was greater than 70%. Jeong (2009) compared title keywords and tags on YouTube and determined that there was a high degree of match between the two sets of terms, specifically more than 50% of terms were shared between metadata fields.

Kipp (2005) compared tags, author keywords, and descriptors using library and information science articles tagged on CiteULike. Many tags were found to be related to both the author keywords and descriptors; however, tags were often not part of the thesauri used by the professional indexers and, thus, were not formally linked to the descriptors. Other terms were identical to thesaurus terms or part of the entry vocabulary of the thesaurus itself (Kipp 2005). Kipp (2011) found that tags were more likely to match author keywords than descriptors (33% of matches were exact matches), although 16% of tag matches to descriptors were exact matches and 19% of author keyword matches to descriptors were also exact matches. The results of both studies suggested that there was sufficient overlap for tags or author keywords to act as entry vocabulary for descriptors or as additional access points to improve retrieval.

A few more recent studies have also compared tagging and controlled vocabularies on academic social bookmarking tools (Lin et al 2006; Kipp 2007b; Bruce 2008; Good and Tennis 2008; Heckner et al 2008; Good and Tennis 2009; Trant 2009). These studies have shown general agreement in their results showing differences between user and professional indexer terminology. Minor differences have been reported between studies, suggesting that comparisons between tagging and controlled vocabularies may be affected by field of study (Kipp 2005; Kipp 2007b). This result matches results from Frost (1989) in which the degree of match between title keywords and subject headings was strongly dependent on the field of study.

This study, therefore, posed the following research questions in order to examine the question of term

248

M. E. I. Kipp. Tagging of Biomedical Articles on CiteULike: A Comparison of User, Author and Professional Indexing

Knowl. Org. 38(2011)No.3

usage and convergence between tags, keywords and descriptors by exploring the tagging phenomenon as it is growing on CiteULike using articles from biomedical journals.

– To what extent do term usage patterns of user tags, author keywords, and professional indexer descriptors suggest that professional indexers are merely engaging in essentially the same activities as authors and users, but merely at a more rigorous, thorough, and consistent level?
– To what extent do term usage patterns suggest that authors and users are engaging in a fundamentally different activity, one that cannot be usefully compared or linked to the activities of professional indexers?

This paper reports on the results of an exploratory study of CiteULike (a social bookmarking service), which compared the tags assigned to academic journal articles by users of the CiteULike bookmarking system to descriptors assigned by professional indexers and to author keywords assigned by authors to their own journal articles.

4.0 Methodology

4.1 Selection of Field of Study, Journals and Descriptors

The prevalence of biological terms in the CiteULike tag cloud at the time of data collection suggested that biology or medicine would be good choices for this study. Journals were selected based on three criteria: relative prominence within the field as defined by the Journal Impact Factor, the presence of author keywords, and the potential for collection of descriptors from an online database. Journals selected for this study were chosen because they are: a) biology related, b) require authors to submit keywords for their articles, and c) are indexed in PubMed using *Medical Subject Headings (MeSH)*. *MeSH* is well known and accepted in the biomedical community, and PubMed is heavily used to locate articles in the medical and biological sciences and thus provides a useful controlled vocabulary for comparison to tags and author keywords. Two academic journals matching the criteria were chosen for this study: *Journal of Molecular Biology* and *Proteins*. These journals were located manually from journal websites and direct examination of sample articles. CiteULike was chosen for this study, as it provides a facility for searching by journal

name, something which is not available in similar tools such as Connotea.

4.2 Selection of Articles

Tag data for this study was collected from CiteULike between January 12, 2007 and January 24, 2007 via a python script (citeulike.py). Data were stored in a MySQL database for further analysis. Data collected from CiteULike (see Figure 1) included article data and post data. The article data consisted of the article title, authors, source (i.e. journal name, volume number, etc.), publication date, abstract (where available), URL, and a list of userids. Post data consisted of an associated articleid, date posted, userid, and a list of tags associated with the article. The data for individual posts associated with each article are stored separately in the researcher's database after retrieval, but linked to the article, so that user tag lists can be compared separately but also combined for comparison with author keywords and descriptors.

All articles from the selected journals, which had been tagged on CiteULike by at least one user, were collected. To ensure that all articles from these journals were collected, the python script was designed to collect articles under all common variants of the chosen journal names (e.g., J. Mol. Biol. for *Journal of Molecular Biology*). These results were parsed to exclude articles which had not yet been tagged by users since CiteULike also provides access to articles from selected journals which have not yet been tagged to assist in the location of new material.

URLs were collected for each article and automatically separated into categories as potential sources of keywords or descriptors. Digital Object Identifiers or DOIs (http://www.doi.org/) were selected by preference as a source of author keywords for journal articles and PubMed URLs were used to locate descriptors (in this case *MeSH* indexing terms).

Author keywords were collected from online journal databases using the DOI (http://www.doi.org/) collected from CiteULike or, in rare cases, by exact title match using Google Scholar.

Professional indexer terms, in the form of descriptors, were located via script access to PubMed (see Figure 2). PubMed provides professional indexer assigned controlled vocabulary subject headers for searchers via *Medical Subject Headings (MeSH)*. Where possible, PubMed URLs and DOI URLs were used directly, as these are often available directly in the CiteULike metadata; otherwise, a script was used to locate PubMed URLs given the DOI, the DOI
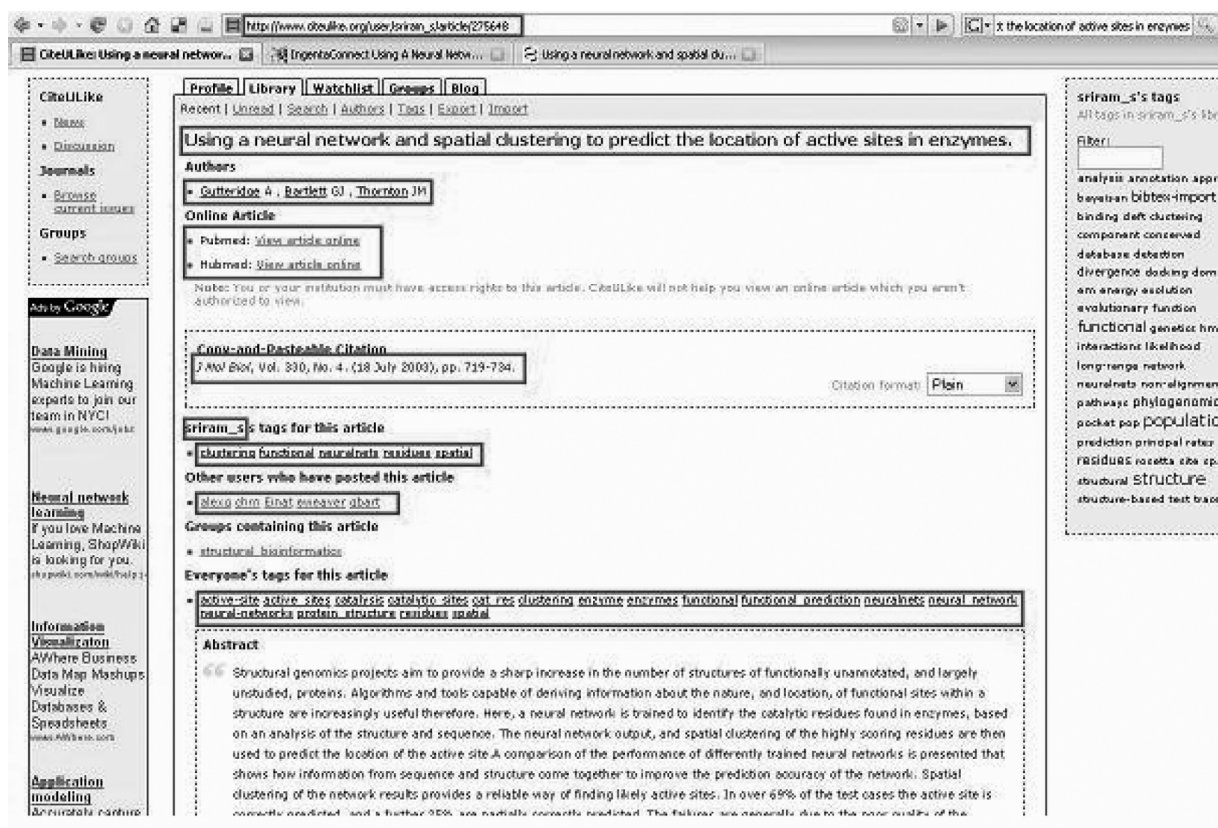
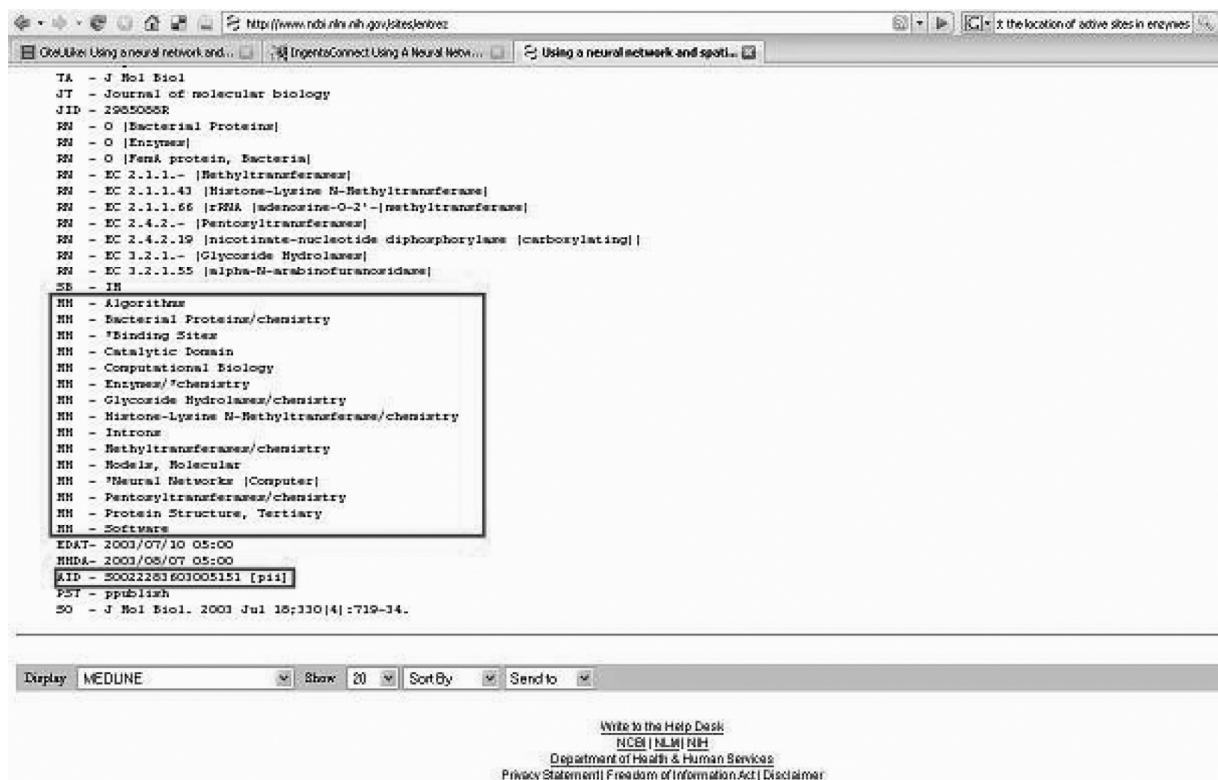*Figure 1.* Sample CiteULike post with collected data highlighted.

Figure 2. Sample PubMed data for an article with MeSH headings highlighted

250

Knowl. Org. 38(2011)No.3

M. E. I. Kipp. Tagging of Biomedical Articles on CiteULike: A Comparison of User, Author and Professional Indexing

given the PubMed ID, or, in extreme cases, Google Scholar was used to locate articles using the article title and other bibliographic information. A total of 19 items could not be located on PubMed and were excluded from the study.

This resulted in a total of 1083 articles for analysis. Since many articles were tagged by more than one user, this resulted in a total of 1588 posts with tag lists for analysis (Table 1).

Journal Name	Number of Articles	Number of Posts
Journal of Molecular Biology	649	931
Proteins	434	657
Total	**1083**	**1588**

Table 1. Journals with author assigned keywords.

4.3 Data Analysis

In the end, each article selected for this study had three sets of terms (tags, author keywords, and descriptors) assigned by three different classes of metadata creators. The data were stored in a MySQL database and preliminary informetrics analysis was done using SQL scripts. Descriptive statistics and basic informetric data were collected to provide a good picture of the scope of the collected data. Additionally, a random sample of articles was selected to have its tags, keywords, and descriptors examined for term usage.

A number of measures of analysis were used including:

– Descriptive statistics (including number of posts per user, number of tags per user, number of tags per article);
– Informetrics methods (especially user vocabulary length and an examination of trends in number of index terms used by professional indexers, authors, and taggers);
– Term comparison; and,
– Thesaural comparison.

Term comparison involved direct examination of terms used by each group and categorisation of terms which did not seem to be directly subject related. Included in this category were methodological terms, geographical terms, proper names, and any other term which was not obviously a subject term.

For the thesaural comparison, user tags, author keywords, and professional indexer-assigned descriptors were compared based on a seven point scale from Kipp (2005). This scale is similar to that used by Voorbij (1998) in a study of title keywords. While Voorbij examined descriptor correspondence to title keywords, this study examines the correspondence between all three sets of tags using a structured thesaurus (*MeSH*) to generate similarity comparisons. Where possible, comparisons have been done across all three sets of terms, but where the term (or any related term) is lacking from one set, the other two sets were compared against the seven categories. Comparisons using this seven category system were done by the author.

The following are the categories as modified:

1. Same: the descriptors and keywords are the same or almost the same (e.g., plurals, spelling variations, acronyms, and multiword terms split into facets);
2. Synonym: the descriptors and keywords are synonyms (corresponds to USED FOR in a thesaurus);
3. Broader Term: the keywords or tags are broader terms of the descriptors in the thesaurus;
4. Narrower Term: the keywords or tags are narrower terms of the descriptors (like Broader Term, this indicates that the user or author term is in the thesaurus as a broader or narrower term of the associated indexer term);
5. Related Term: the keywords or tags are related terms of the descriptors;
6. Related Not In Thesaurus: there is a relationship (conceptual, etc.), but it is not obvious to which category it belongs or it is not formally in the thesaurus;
7. Not Related: the keywords and tags have no apparent relationship to the descriptors, also used if the descriptors are not represented at all in the keyword and tag lists (Kipp 2005).

5.0 Results

5.1 Authors, Users and Journals

Bibliographic data for a total of 1083 articles were collected from CiteULike. This data set included all articles tagged by at least one user from the chosen journals: Proteins and Journal of Molecular Biology. The data set thus contained a total of 1,588 posts.

Unique user names present in the sample totaled 239. Since it is possible for a user to create a second

Knowl. Org. 38(2011)No.3
M. E. I. Kipp. Tagging of Biomedical Articles on CiteULike: A Comparison of User, Author and Professional Indexing

251

account with a different email address, it is not possible to ensure that these are indeed 239 distinct persons. Each user name was associated with at least one post in the data set. One user had posted 94 of the 1,588 collected posts. Many other users had posted significantly fewer posts (Top 5 users posted 94, 65, 64, 44, and 43 posts, respectively). A total of 94 users (39%) had posted only one post in the data set. Of the users who posted more frequently in this data set, 42 (18%) posted 10 or more times.

A similar drop off can be seen in the data set when examined based on the number of users who have posted a link to a specific article. In this case, the maximum number of users per article was 14, the minimum 1, and the median 2 (Table 2).

Number of Users per Article	Article Title
14	Principles of docking: An overview of search algorithms and a guide to scoring functions.
7	Comparing protein-ligand docking programs is difficult.
6	Protein flexibility predictions using graph theory.
6	Binding MOAD (Mother Of All Databases).
6	The Relationship between the Flexibility of Proteins and their Conformational States on Forming Protein-Protein Complexes with an Application to Protein-Protein Docking

Table 2. Number of users who posted a link to a specific article.

In fact, the number of users who posted more than one article dropped off quite quickly (799 articles were posted only once, median was 1 post per article). This shows similarities to findings from citation analysis, which shows that a few articles tend to be highly cited while many others are infrequently cited (Table 3). Citation analysis shows that a power law occurs in citations and this study shows that a power law also exists in posting of articles to CiteULike. This suggests another difference from professional indexing in which professionals will index according to their own consistent and exhaustive policies while users may stop indexing after only a few posts.

Number of Articles	Number of Users Who Posted X articles
1	799
2	195
3	64
4	25
5	10
6	6
7	1
8	0
9	0
>=10	1

Table 3. Number of users who posted X articles.

5.2 Tags, Keywords and Descriptors

The total number of descriptors in the sample was found to be extremely high. This is due to the fact that PubMed articles tend to have many descriptors assigned to increase recall and to cover categories such as methodology of the study and user groups studied (Table 4).

	Tags	Keywords	Descriptors
Unique	1136	3181	2746
Total	3788	4866	12473

Table 4. Number of indexing terms of each type.

Many tags, keywords, and descriptors occurred frequently in the collected data. The most popular tag was "protein_structure," used 140 times; the most popular keyword was "protein folding," used 58 times; and, the most popular descriptor was "Models, Molecular," used 649 times in the data set (Table 5).

Frequency	Tag
140	protein_structure
114	no-tag
114	protein
103	structure
97	docking

Table 5. Most commonly used tags.

252

M. E. I. Kipp. Tagging of Biomedical Articles on CiteULike: A Comparison of User, Author and Professional Indexing

Knowl. Org. 38(2011)No.3

A total of 645 tags were used only once in the data set, and 185 tags were only used twice. The median number of times a tag was used in the data set was 1.

In comparison, author keywords were much more diverse with 2,548 of the keywords being used only once in the data set. The maximum number of times a keyword was used was 58, minimum 1, and median 1 (Table 6).

Frequency	Author Keywords
58	protein folding
49	protein structure
46	molecular dynamics
38	protein structure prediction
31	docking

Table 6. Most commonly used author keywords.

Descriptors were heavily reused in the data set, with some descriptors being used hundreds of times. The maximum number of times a descriptor was used in the data set was 649, minimum 1, and median 2 (Table 7).

Frequency	Descriptors
649	Models, Molecular
511	Protein Conformation
388	Proteins
306	Amino Acid Sequence
280	Binding Sites

Table 7. Most commonly used descriptors.

Out of a total of 2746 unique descriptors, 731 descriptors were used only once, and 249 were only used twice. This is a higher reuse rate than that for author keywords.

When examined at the article level, there are similar patterns of usage of tags, keywords, and descriptors. While some articles were highly tagged, the majority had only a few tags. The maximum number of tags assigned to an article was 29, minimum 1, and median 2. The article with 29 tags was tagged by 14 users, suggesting that this is still an example of users assigning some 1-3 tags to an article (Table 8).

An examination of the number of tags per post (an article may be posted multiple times thus generating multiple posts per article) shows smaller numbers of tags. The maximum number of tags per post was 15, minimum 1, and median 2.

Frequency	Article Title
29	Principles of docking: An overview of search algorithms and a guide to scoring functions.
20	Binding MOAD (Mother Of All Databases).
19	Universally conserved positions in protein folds: reading evolutionary signals about stability, folding kinetics and function.
18	How different amino acid sequences determine similar protein structures: The structure and evolutionary dynamics of the globins
18	Using a neural network and spatial clustering to predict the location of active sites in enzymes.

Table 8. Number of Tags per Article (top 5).

Frequency	Article Title
13	Automated prediction of domain boundaries in CASP6 targets using Ginzu and Rosetta-DOM.
13	Automated prediction of CASP-5 structures using the Robetta server.
11	Structure modeling, ligand binding, and binding affinity calculation (LR-MM-PBSA) of human heparanase for inhibition and drug design.
11	Discrimination between native and intentionally misfolded conformations of proteins: ES/IS, a new method for calculating conformational free energy that uses both dynamics simulations with an explicit solvent and an implicit solvent continuum model
10	Minimizing false positives in kinase virtual screens.

Table 9. Number of Keywords per Article (top 5).

Similarly, the maximum number of keywords found for an article in the data set was 13, minimum 1, median 5. One reason why the median number of keywords is higher than for tags is due to the fact that many journals have a set number of author keywords they request, often 5 or 6 (Table 9).

The total number of descriptors used in the data set was 12,743, but the number of unique descriptors was only 2,746. An examination of the number of descriptors per article shows that many articles had a much larger number of assigned descriptors than either tags or keywords. The maximum number of descriptors assigned was 36, minimum 2, median 11. This high

median suggests that PubMed indexers attempt to provide as broad a list of relevant descriptors as possible to aid in information retrieval (Tables 10-11).

Frequency	Article Title
36	Crystal structure of cone arrestin at 2.3A: evolution of receptor specificity.
30	G-protein-coupled receptor domain overexpression in Halobacterium salinarum: Long-range transmembrane interactions in hepta-helical membrane proteins.
29	A Snapshot of Viral Evolution from Genome Analysis of the Tectiviridae Family.
28	Computer-assisted identification of cell cycle-related genes: new targets for E2F transcription factors,
27	Catalytic Independent Functions of a Protein Kinase as Revealed by a Kinase-dead Mutant: Study of the Lys72His Mutant of cAMP-dependent Kinase

Table 10. Number of Descriptors per Article (top 5).

Number of Index Terms (Tags, Keywords or Descriptors) assigned to an article	Tags	Keywords	Descriptors
1	29	0	0
2	20	1	0
3	18	10	0
4	12	16	3
5	4	60	7
6	4	10	4
7	3	1	7
8	3	1	9
9	2	0	11
10	0	0	8
11	1	0	7
12	1	0	9
13	0	1	8
14	2	0	2
15	0	0	4
16	0	0	4

Number of Index Terms (Tags, Keywords or Descriptors) assigned to an article	Tags	Keywords	Descriptors
17	1	0	5
18	0	0	2
19	0	0	2
20	0	0	1
21	0	0	3
22	0	0	2
23	0	0	1
24	0	0	0
25	0	0	1

Table 11. Number of tags, keywords and descriptors applied to individual articles. Each number in the table represents the total number of articles with 1, 2, 3... 25 index terms assigned (number of index terms is the total number of unique terms).

The correlation value obtained when comparing authors versus keywords, again, did not show a significant relationship. This is reasonable as journals request a certain number of keywords per article and thus there is unlikely to be a relationship between the number of keywords and the number of authors. The correlation value for users versus tags did show a significant relationship with an R^2 value of 0.619 ($p < 0.05$). The correlation value for users versus unique tags also showed a significant relationship with an R^2 value of 0.563 ($p < 0.05$). These results suggest that there is a significant positive correlation between the number of users and the total number of tags (unique or not) assigned to an article. This result is significant for this data set, but similar results were found in Kipp (2005) and Kipp (2011) with a different data set.

The largest user vocabulary length in the data set was 62, the smallest 1, and the median 2. This suggests that most users tend to use a small number of tags, while a small number of users will use more tags. When the user vocabulary length is broken down at the individual article level, the largest length was 15 tags for one article (Table 12).

254

Knowl. Org. 38(2011)No.3

M. E. I. Kipp. Tagging of Biomedical Articles on CiteULike: A Comparison of User, Author and Professional Indexing

User	Max tag list length	Min tag list length	Number of articles posted
3109	7	2	15
3063	6	1	73
4068	15	2	9

Table 12. User Vocabulary Length by Article.

5.3 Term Usage

Examining the tags from a specific article (788), "Computer modeling 16 S ribosomal RNA," it was noted that nine tags were applied to the article (Table 13). Two of the tags came directly from the title, namely 'rna' and '16s'. It is interesting that taggers chose to use the term "algorithms" rather than a term like "computer modeling," which was used for other items in the data set, despite the fact that computer modeling is a term from the title. In fact "computer modeling" is one of the author keywords for this article and the term "computer simulation" occurs in the descriptor list. It is worth noting here that, while social bookmarking systems like delicious.com offer lists of suggested terms for tagging, CiteULike does not offer any prompting to users to aid them in selecting tags.

Additional terms that do not come directly from the title were 3d, prediction, distance_geometry, bioinformatics, structure, and structure_prediction. The term bioinformatics is an excellent example of an extremely generic term for computer modelling and analysis as related to biology, which one would not necessarily expect in the descriptor list since it would likely be a Broader Term. Seen across all three sets of indexing terms are variants on "16s rna."

Tags	Keywords	Descriptors
3d	16 S RNA	Base Sequence
algorithms	ribosome	Computer Simulation
prediction	computer modeling	Cross-Linking Reagents
rna	distance geometry	Escherichia coli
16s		Models, Molecular
distance_ geometry		Molecular Sequence Data
bioinformatics		Nucleic Acid Conformation

Tags	Keywords	Descriptors
structure		RNA, Ribosomal, 16S
structure_ prediction		

Table 13. Tags, Keywords and Descriptors for Article 788.

5.4 Thesaural Relations

For the thesaural analysis of the second data set, a random sample of 500 articles was selected to be analysed. Again, the most common relationships were Equal, Related Term, and Related but not in the thesaurus. Unlike the LIS data set, however, Related Term was more common than Equal. This may be due to the extensive entry terminology and increased number of related terms in the *MeSH* thesaurus or that there is a substantial vocabulary for this knowledge area, including specialist and nonspecialist terms, making it less likely that the three indexer groups will converge on the exact same term.

It is worth noting that the prevalence of non-matching terms does not indicate that these terms are irrelevant to the article. Most non-matching terms were actually completely topical, but were simply not used by more than one of the indexing groups. Of the 707 non-matches in the sample, 23 or 3% were judged to be Not Related to the subject of the article. This is much higher than the expected incidence of Not Related terms in a standard bibliographic database, but is to be expected in a database with user tags since terms which are not subject related have been shown to be popular when users organise material (Malone 1983; Kwasnik 1991; Kipp 2007a).

Using the modified version of Voorbij's scale, it was found that the most common relationship discovered in the groups of user, author and professional indexer keywords examined was category 6 or Related Not In Thesaurus. This form of relationship occurred in 65 of 100 articles or 65%. The next most common relationship was the Related Term (RT) relationship at 64%, followed by Same with 48%. This is a slight reversal of the findings in Kipp (2005) where the Same relationship was more common than the RT relationship. Following this was Synonym in 32 articles and Narrower Term and Broader Term combined in 14 articles. Not Related terms occurred in 91% of cases. On average 3 Not Related terms occurred per article (Table 14).

In total, there were 707 Not Related terms and 473 matches in the thesaural comparisons. Related Term

	Same	Synonym	NT/BT	RT	Related	Not Related
0	52	68	86	36	35	9
1	19	26	8	26	21	4
2	15	3	5	14	17	4
3	13	3	1	9	11	8
4	1	0	0	10	9	6
5	0	0	0	4	1	7
6	0	0	0	0	4	10
7	0	0	0	2	0	6
8	0	0	0	0	0	7
9	0	0	0	0	0	10
10	0	0	0	0	0	6
>10	0	0	0	0	1	23
Total Matches (1- >10)	48	32	14	64	65	91
Sum by Frequency of Matches (1-10)	92	41	21	155	164	707

Table 14. Frequency of occurrence of the thesaural comparison categories. The left column represents the number of articles with 0, 1,2 ... matches of that type. Each number in the table represents the total number of matches (either binary or trinary) between the three sets of index terms. Note that the sum of matches represents the sum of all matches not the sum of the frequencies. This value is calculated by adding the totals multiplied by the frequency.

(RT in a thesaurus) 155 matches and Same (identical to the descriptor) at 92 matches were the most common of the thesaural comparisons. There were 164 terms that were Related Not In Thesaurus. This, and the high number of non-matches, suggests that while users often use terminology which is somewhat like that used in a thesaurus, they tend not to use the exact terminology of the thesaurus to describe their work. This tends to reinforce the idea that tagging could be very useful in providing an entry vocabulary to the traditional controlled vocabulary, allowing users the benefits of both systems.

Though thesaural relations were less common, many matches did fall into the Same or Related term categories, and some 20% of articles had Narrow Term/Broader Term or Synonym matches as well. These relationships were less common than the final two non-thesaural categories, covering the Related Not In Thesaurus and Not Related categories respectively. In total, the thesaural relations accounted for 309 matches out of 473 total matches or 65% of all matches. This includes the Same (equivalence) category, Synonyms, Broader Terms, Narrower Terms, and Related Terms.

Binary comparisons were more common than trinary comparisons. In total, there were 392 binary

matches versus 81 trinary matches. The most common trinary relationship was Related Not In Thesaurus, as might be expected. This was also the most common binary relationship (Table 15).

	Binary Matches	Trinary Matches	Total Matches
Same	78	14	92
Synonym	33	8	41
Narrower or Broader Term	17	4	21
Related Term	129	26	155
Related	135	29	164

Table 15. Comparison of binary versus trinary matches.

The number of comparisons per article was somewhat dependent on the length of the term lists for tags, keywords and descriptors. An article with a higher number of tags, keywords, and descriptors would have a higher chance of having a larger number of matches and would also likely have more non-matches.

The maximum number of occurrences of specific matches shows, again, that binary matches are generally more common than trinary matches. The maxi-

256

Knowl. Org. 38(2011)No.3

M. E. I. Kipp. Tagging of Biomedical Articles on CiteULike: A Comparison of User, Author and Professional Indexing

mum number of matches of any kind per article was 12, the minimum 2, and the median 3 (Table 16).

	Binary Matches	Trinary Matches
Same	4	1
Synonym	3	2
Narrower or Broader Term	2	3
Related Term	7	4
Related	5	7
Not Related		22

Table 16. Maximum number of occurrences of each match per article.

Trinary matches involved an index term from each of the three user categories; binary matches only involved terms from two of three categories. While author/professional indexer matches were most common overall, when normalised, it proved to be author/user matches in the Same category that were the most common of the matches. Author/user matches were more likely to be thesaural matches while author/professional indexer matches were less likely to be thesaural matches (Table 17). One potential limitation of this study is that it is impossible to ensure that items tagged by only one person have not been tagged by the article author. Since author/users matches are the most common category of thesaural matches, there remains a possibility that users tagging articles may in some cases actually be the authors of the articles in question. This becomes an issue since authors may have an incentive to promote their articles on CiteULike, an issue which would not occur in a traditional journal database. However, it remains impossible to match a CiteULike user name to the name of an author of an article.

5.5 Related Tags

Many relationships fell into the 6th category (35%)— Related Not In Thesaurus. This category included relationships that were ambiguous or difficult to fit into categories 1-5, as well as relationships that were not formally listed in the thesaurus but suggested by user tags, author keywords, or PubMed's entry vocabulary. Common relationships included: the relationship between an object and its field of study, the relationship between two fields of study which examine different aspects of the same phenomenon, and the use of a methodology or form of inquiry in a new environment.

Examples of Related Terms include "structure" (a user tag) and "Models, Molecular" (a *MeSH* heading from the thesaurus). The link is suggested by entry vocabulary under "Models, Molecular." This was a very common relationship in the sample, as users chose to use less specific terminology, perhaps because there is a tacit assumption that the article is related to molecules, or proteins, or some other area of study and therefore it is not necessary to add these terms to their own tag lists. Other examples of Related Terms are the author keyword (and occasionally also user tag) protein-families. This term is related to Proteins in the *MeSH* thesaurus, but is not listed as entry vocabulary. The term is used to refer to relationships between proteins which are not yet included in the thesaurus. Another example of a set of Related Terms is the relationship between the author keyword "thermal unfolding" and the descriptor "Protein Denaturation." When proteins are heated (subjected to thermal stresses), they break down or denature.

Newer terminology or highly specific terminology for newly discovered structures is created as discoveries are made. Examples of this type of term were found in both user and author terminology. For ex-

	User/Professional		Author/Professional		Author/User	
	Raw	Percent	Raw	Percent	Raw	Percent
Same	16	0.16	40	0.16	22	0.48
Synonym	11	0.11	19	0.08	3	0.07
Narrower or Broader Term	4	0.04	13	0.05	2	0.04
Related Term	41	0.4	80	0.33	8	0.17
Related Not in Thesaurus	30	0.29	94	0.38	11	0.24
Totals	102	1	246	1	46	1

Table 17. Comparison of number of binary matches between user/professional, author/professional and author/user.

ample, the acronym "PISEMA" for Polarization Inversion Spin Exchange at Magic Angle was an author keyword related to the descriptor term "Magnetic Resonance Spectroscopy." Another example was the author term "PISA wheels" for a particular form of secondary protein structure, which was represented by the descriptor "Protein Structure, Secondary."

This inclusion of newer terms in the user tags can happen faster than it would in a traditional thesaurus or other controlled vocabulary, as one of the goals of a thesaurus is to reproduce the accepted state of knowledge in a field, which leaves the leading edge of the field time to determine standard terminology that will eventually be added to the thesaurus.

5.6 Unrelated Tags

Tags, keywords, and descriptors falling into the 7th category (Not Related) tended to fall into six basic types: time and task management, geographic or personal, specific details and qualifiers, generalities, emergent vocabulary, and other. Since the author of this paper does not want to presume that the thesaurus is inherently superior in its indexing, descriptors that did not match any terms used by the author or users were also placed in this category.

Of the Not Related terms (52 tags, 89 author keywords, and 543 descriptors), the majority were subject related but were simply not used by any of the other two indexing groups. Many of these terms were descriptors which the authors or users simply did not use. The large number of subject applicable descriptors which are not matched by tags and author keywords indicates that descriptors continue to provide an important contribution to indexing even if author keywords and user tags are included in the mix.

Many terms, especially descriptors, were frequently not matching. Terms such as Models, Molecular would fit into the category of generalities, since the term is a quite general term for molecular modeling (Table 18).

Terms	Frequency
Models, Molecular	26
Animals	20
Molecular Sequence Data	18
Amino Acid Sequence	18
Binding Sites	13

Table 18. Most common descriptors which did not match author or user terms.

Time and task related terms did not occur in the descriptors at all and were rare in the author keywords. Most time and task related terms were tags. One tag was a specific date '31mar06' while others appeared to be references to projects or groups "cafasp," "OR-Fans," "refs_ox." An author keyword used to tag one article was "drug design" describing the project or purpose of the research. This term was not echoed in the descriptors. No geographic terms were present in the sample, but two tags that were related to the authors of a paper were located: "jwm_author" and "prossnitz." As well, one non-matching tag was found to be the name of a specific pharmaceuticals company from the UK: "inpharmatica."

Many examples of specific details and qualifiers were found in the sample. This area is one of the areas in which users, authors, and professional indexers often appear to disagree on which aspects of a paper are most important.

Methodology terms were common in all three groups. User terms like "bioinformatics" were used instead of the *MeSH* descriptor "Computational Biology." Terms that matched on one article would be missing a match on other articles because users or authors did not consider the precise methodology to be an important enough aspect of the work to index. Other examples of methodologies include the descriptor "Crystallography, X-Ray," the author keyword "Smith-Waterman," and the tag "phi_value_analysis."

Another important group in this category is user groups, in other words the group being studied. This group was also present in Kipp (2005). However, user group terms were almost entirely descriptors. Examples included "Animals," "Humans," and "Leopard Frog." This finding is similar to previous studies involving academic journals (Kipp 2005), but is distinct from a study involving articles from *JAMA* (a professional journal) in which user group terms were more prevalent in the user tags (Kipp 2007b).

Many descriptors fell into the category of generalities. Terms such as "Models, Molecular," "Models, Theoretical," "Models, Chemical," and "Models, Biological" are all general descriptors discussing methodologies or domain specific modeling techniques. One user tag fell into this category as well: "mathematical_model," which was not matched by a similar descriptor, suggesting that while the user was interested in the modeling techniques in the article they were not deemed important enough to be listed in the descriptors.

Also present were terms that constituted emergent vocabulary. One common example of emergent vo-

258

Knowl. Org. 38(2011)No.3

M. E. I. Kipp. Tagging of Biomedical Articles on CiteULike: A Comparison of User, Author and Professional Indexing

cabulary was the term bioinformatics. This term represents the melding of computer science/data mining and biology. Although terms do exist in *MeSH* for this field, notably "Computational Biology," the term bioinformatics is most commonly used by users. Another common term was "protein families" used to describe related proteins. This term is also not present in *MeSH* although it was used by multiple users.

A small set of terms, 22 in total, did not appear to be related to the subject of the article. These non-subject tags have been reported in other studies (Kipp and Campbell 2006; Kipp 2007a) and were generally time and task related. The system assigned tag "no-tag" was the most common and occurred five times in the sample. A similar tag "to-be-tagged" occurred 3 times in the sample. Although affective terms were present in the full data set, none were found in the sample.

6. Discussion and Conclusions

As previously discussed in Kipp (2005), Kipp and Campbell (2006), and Hammond et al (2005), users use some terminology which is rare or completely absent from author keyword lists or descriptor lists. Time and task related terminology were present in the current study, as well as earlier studies of academic social bookmarking tools (Kipp 2005; Kipp 2007b; Heckner et al 2008). Terms such as "to_be_tagged," "toread," and a number of calendar dates (e.g., 31/03/06, 1998) were found as tags assigned to articles in this data set.

While professional indexers considered geographic location to be an important part of the description of the aboutness of an article, authors, and users tended to assume it was somewhat less important than the other contexts of the articles. In many cases, this may be true. For example, the difference between an information retrieval study performed in the United Kingdom and one performed in the United States is probably not significant due solely to the difference in geographic location.

Many user terms were found to be Related (Not In Thesaurus) to the author and professional indexer terms, but were not part of the formal thesaurus used by the professional indexers and, thus, not formally linked to the professional indexer terms. In some cases, this was due to splitting of multiword terms for example "protein" and "structures" used separately in the tag lists where they were linked in the thesaurus or the use of abbreviations such as "PDB" for "Databases, Protein." In some cases, this was due to the use

of broad terms that were not included in the thesaurus such as information, knowledge, or computers. Heckner et al (2008) also report that many user terms are more general than author or professional terms. In some cases, this was also due to the use of newer terminology (web2.0, folksomonies, tagging) or to differences in approach to a problem (information seeking versus information retrieval).

Terms such as "human" and "animal" showed that users tagging biology related articles are extremely interested in methodology and user groups associated with articles. This is distinct from Kipp (2005) where such terms were more common in the descriptors unless they described extremely specific kinds of methodologies, such as "pubmed-mining" for data-mining of PubMed. Additionally, taggers assigning tags to academic articles have some specific terminology requirements such as methodology or user group being studied, which are not present in the same quantity in studies examining more free form sites such as delicious.com (Kipp and Campbell 2006).

The differing terminologies of various users groups, frequently discussed in the indexing literature in the guise of entry vocabulary, has also been noted in additional studies of tagging terminology. Trant (2009) examined the terminology used by museum visitors versus museum cataloguers and determined that both sets of terminology provided usable but different views of the object. While users provided terms describing what they saw, cataloguers provided description appropriate to the provenance of the item. Both sets were useful, but came from a different tradition, and the study suggests both sets of terms would be useful for search and discovery (Trant 2009). Heckner et al (2008) found that 46% of their collected tags were not directly from the text suggesting that user tags are indeed adding "to the lexical space of the tagged resource" (Heckner et al 2008).

This study has implications for the design of systems for accessing, indexing, and searching document spaces. The popularity of Google has demonstrated that users prefer to be able to search for items in a more natural way using one interface to locate items of a varied nature. However, users also express frustration at being unable to locate items or narrow their search results from a huge search set, for example 300,000 hits on Google (Campbell and Fast 2004). Controlled vocabularies help to narrow a search set to a manageable size, but controlled vocabulary usage can be expensive and may require user training for effective search.

Morville (2005) suggests that the beauty of metadata such as controlled vocabularies is that they and tags are not mutually exclusive. He links tagging and controlled vocabularies using Stewart Brand's concept of pace layering. Pace layering refers to the process of constructing a building from the physical building, which changes slowly, to the office supplies and companies which use the building, which change more quickly. Morville suggests that tags will gain their greatest utility as a fast layer on top of the slower layer of controlled vocabularies. (Morville 2005, 140-1) This study's findings, specifically that 16% of tag matches to descriptors were in the Same category, 11% to Synonyms (often entry vocabulary), 40% to Related Terms, and 29% to Related Not in Thesaurus, suggest that there are enough commonalities to make links between the fast layer of tags and the slow layer of controlled vocabularies as well as sufficient differences to make it worthwhile to maintain the two separate layers.

Of the Not Related terms, which included terms that were task or project oriented, the majority were subject-related, but only used by one of the three groups of indexers (professional indexers, authors, or taggers). In fact, only 3% of the Not Related terms were judged to be unrelated to the subject of the article. Many descriptors ended up as Not Related terms despite being topically relevant because authors or taggers were less comprehensive in their indexing. The presence of such a large number of subject applicable descriptors that were not matched by all three groups of indexers suggests rather strongly that descriptors applied by a professional indexer continue to provide an important contribution in the indexing of journal articles by providing comprehensive subject access, while the author keywords and user tags provide a potentially more holistic view of the subject relevance of the article including input from researchers from a variety of related fields. The presence of task oriented tags such as @toread indicates that users are blending subject and associated indexing into the system in ways which were not supported by separate OPACs and reference management tools but are supported by social bookmarking tools like CiteULike or newer OPAC interfaces that support social tagging and user annotations.

While a majority of tags (and author keywords) were subject related and many matched terms in the thesaurus, it is important to recognise that many of these matches were to Related Terms and not to the specific descriptors chosen by professional indexers. Additionally, some matches were made to entry vo-cabulary (Synonyms) and others to the category of Related Not in Thesaurus. The prevalence of matches to terms which are not descriptors or not yet descriptors in the case of emergent terminology, has implications for search, since it highlights the issue of differences between professional indexing vocabulary and user vocabulary. While the issue of user vocabularies has been studied previously, this has not generally resulted in changes to OPACs or article databases.

Weinberger (2007) notes that, contrary to the hierarchical tree-like structures of a controlled vocabulary, tagging is more like a pile of leaves with all tags appearing at different levels of specificity from very general to very specific. The mixture of these terms in a flat folksonomy is a far cry from the traditional hierarchical system, but still provides some measure of access at different specificity levels. Rather than attempting to remove or hide the ambiguity, tags display it all and allow the user to select the appropriate level of specificity or generality. (Weinberger 2007, 93-95) Additionally, the tags provide a different kind of organisational system from that found in traditional organisational systems, since they include connections to the user who provided the metadata as well as the metadata itself. This personal connection is distinctly different from traditional systems.

The differing terminology used in tag lists suggests that tagging may be a working example of Vannevar Bush's associative trails. He argued that associative trails better represented how users actually work with their documents: by association rather than by categorisation (Bush 1945). This suggests that user tagging could provide additional access points to traditional controlled vocabularies and provide users with the associative classifications necessary to tie documents and articles to time and task relationships as well as other associations which are new and novel.

Studies showing that author keywords, title keywords, and tags provide additional useful terms for search and information retrieval suggest that systems should begin to include these terms in the metadata and provide users with the ability to filter, cluster, sort, search and organise using subject terms assigned by professional indexers, keywords assigned by authors and tags assigned by themselves or other users. At minimum, data from author keywords, title keywords, and tags can be invaluable data for updates to indexing systems in terms of entry vocabulary and emergent terminology.

This study demonstrated that, while many tags and author keywords are equivalent to descriptors, others add additional information, both classificatory and

260   Knowl. Org. 38(2011)No.3

M. E. I. Kipp. Tagging of Biomedical Articles on CiteULike: A Comparison of User, Author and Professional Indexing

associative, which may be beneficial to users. In addition, the presence of descriptors that are not matched by tags or author keywords demonstrates that descriptors continue to perform a useful function in indexing articles, even when tagging is present. Findings from this study demonstrates that traditional systems need to make better use of their existing indexing languages by allowing users access to broader, narrower and related terms, perhaps especially related terms since matches with tags are so prevalent, when searching and browsing capitalising on the significant investment in subject indexing. Additionally, traditional systems can be enhanced by contributions from article authors and taggers who provide a more expansive picture of the relevance and contributions of articles to a field of knowledge and to related fields. While the collection of author keywords and tags for supplementing descriptors would have been expensive in the past, the increasing move to electronic journal articles can actually be beneficial for subject indexing as it provides access to more information such as author keywords and tags which can enhance the process of knowledge discovery.

References

Ansari, Mariam. 2005. Matching between assigned descriptors and title keywords in medical theses. *Library review* 54: 410-14.

Bloomfield, Masse. 1966. Simulated machine indexing, Part 2: Use of words from title and abstract for matching thesauri headings. *Special libraries* 57: 232-35.

Bruce, Robert. 2008. Descriptor and folksonomy concurrence in education related scholarly research. *Webology* 5n3. http://www.webology.ir/2008/v5n3/a59.html (accessed June 24, 2010).

Bush, Vannevar. 1945. As we may think. *The atlantic monthly* 17n: 101-08. http://www.theatlantic.com/doc/194507/bush (accessed June 24, 2010).

Campbell, D. Grant and Fast, Karl V. 2004. Panizzi, Lubetzky, and Google: How the modern web environment is reinventing the theory of cataloguing. *The Canadian journal of information and library science* 28: 25-38.

Chan, Lois Mai. 1989. Inter-indexer consistency in subject cataloging. *Information technology and libraries* 8: 349-58.

Cleveland, Donald B., and Cleveland, Ana D.. 1983. *Introduction to indexing and abstracting*. Littleton, Col.: Libraries Unlimited Inc.

Eisenstein, Elizabeth L. 1983. *The printing revolution in early modern Europe*. Cambridge: Cambridge University Press.

Frost, Carolyn. 1989. Title words as entry vocabulary to *LCSH*—correlation between assigned *LCSH* terms and derived terms from titles in bibliographic records with implications for subject access in online catalogs. *Cataloging & classification quarterly* 10 no. 1: 165-79.

Gil-Leiva, Isidoro, and Alonso-Arroyo, Adolfo. 2007. Keywords given by authors of scientific articles in database descriptors. *Journal of the American Society for Information Science and Technology* 58: 1175-87.

Good, Benjamin M. and Tennis, Joseph T. 2008. Evidence of term-structure differences among folksonomies and controlled indexing languages. *Proceedings of the Annual Meeting of the American Society for Information Science and Technology*, Columbus, Ohio, USA, October 24-29, 2008. http://www.asis.org/Conferences/AM08/posters/78.html (accessed June 24, 2010).

Good, Benjamin M., and Tennis, Joseph T. 2009. Term based comparison metrics for controlled and uncontrolled indexing languages. *Information research* 14n1: paper 395. http://informationr.net/ir/14-1/paper395.html (accessed June 24, 2010).

Hammond, Tony, Hannay, Timo, Lund, Ben, and Scott, Joanna. 2005. Social bookmarking tools (I): A general review. *D-lib magazine* 11n4. http://www.dlib.org/dlib/april05/hammond/04hammond.html (accessed June 24, 2010).

Heckner, Markus, Mühlbacher, Susanne, and Wolff, Christian. 2008. Tagging tagging. Analysing user keywords in scientific bibliography management systems. *Journal of digital information* 9n2. http://journals.tdl.org/jodi/article/view/246 (accessed June 24, 2010).

Jeong, Wooseob. 2009. Is tagging effective?—overlapping ratios with other metadata fields. In International Conference on Dublin Core and Metadata Applications, Seoul, South Korea. DC-2009 Seoul Proceedings. http://dcpapers.dublincore.org/ojs/pubs/article/view/975 (accessed June 24, 2010).

Kipp, Margaret E.I. 2005. Complementary or discrete contexts in on-line indexing: A comparison of user, creator and intermediary keywords. *Canadian journal of information and library science* 29: 419-36. http://eprints.rclis.org/8379/ (accessed June 24, 2010).

Kipp, Margaret E. I. 2007a. @toread and cool: Tagging for time, task and emotion. *Proceedings of the*

8th Information Architecture Summit, Las Vegas, USA, March 22-26. http://eprints.rclis.org/10445/ (accessed June 24, 2010).

Kipp, Margaret E.I. 2007b. Tagging for health information organisation and retrieval. *North American Symposium on Knowledge Organisation (NASKO), Toronto, Ontario, Canada, June 14-15, 2007.* http://eprints.rclis.org/11412/ (accessed June 24, 2010).

Kipp, Margaret E.I. 2011. User, author and professional indexing in context: An exploration of tagging practices on CiteULike. *Canadian journal of library and information science* 35n1 (forthcoming).

Kipp, Margaret E.I., and Campbell, D. Grant. 2006. Patterns and inconsistencies in collaborative tagging practices: An examination of tagging practices. *Annual General Meeting of the American Society for Information Science and Technology, Austin, TX, USA, November 3-8, 2006.* http://eprints.rclis.org/archive/00008315/ (accessed June 24, 2010).

Kwasnik, Barbara H. 1991. The importance of factors that are not document attributes in the organisation of personal documents. *Journal of documentation* 47: 389-98.

Lin, Xia, Beaudoin, Joan E., Bui, Yen, and Desai, Kaushal. 2006. Exploring characteristics of social classification. *Proceedings of the 17th Workshop of the American Society for Information Science and Technology Special Interest Group in Classification Research*, pp. 1-19.

Malone, Thomas W. 1983. How do people organise their desks? Implications for the design of office information systems. *ACM transactions on office information systems* 1: 99-112.

Markey, Karen. 1984. Interindexer consistency tests: A literature review and report of a test of consistency in indexing visual materials. *Library & information science research* 6: 155-77.

Mathes, Adam. 2004. Folksonomies - Cooperative classification and communication through shared metadata. *Adammathes.com.* http://www.adammathes.com/academic/computer-mediated-communication/folksonomies.html (accessed June 24, 2010).

Morville, Peter. 2005. *Ambient findability: What we find changes who we become.* Sebastopol, CA: O'Reilly.

Schultz, Claire K., Schultz, Wallace L., and Orr, Richard H. 1965. Comparative indexing: Terms supplied by biomedical authors and by document titles. *American documentation* 16: 299-312.

Shirky, Clay. 2005. Ontology is overrated: Categories, links, and tags. *Shirky.com.* http://shirky.com/writings/ontology_overrated.html (accessed June 24, 2010).

Strader, C. Rockelle. 2009. Author-assigned keywords versus Library of Congress Subject Headings: Implications for the cataloging of electronic theses and dissertations. *Library resources & technical services* 53: 243-50.

Trant, Jennifer. 2009. Tagging, folksonomy and art museums: Results of steve.museum's research. Technical Report. http://conference.archimuse.com/jtrants/stevemuseum_research_report_available (accessed June 24, 2010).

Voorbij, Henk J. 1998. Title keywords and subject descriptors: A comparison of subject search entries of books in the humanities and social sciences. *Journal of documentation* 54: 466-76.

Weinberger, David. 2007. *Everything is miscellaneous: The power of the new digital disorder.* New York, NY: Times Books.