COMMENTARY

D-Lib Magazine
January 2006

Volume 12 Number 1

ISSN 1082-9873

Folksonomies

Tidying up Tags?

Marieke Guy
UKOLN
<m.guy@ukoln.ac.uk>

Emma Tonkin
UKOLN
<e.tonkin@ukoln.ac.uk>

1. Introduction

A folksonomy is a type of distributed classification system. It is usually created by a group of individuals, typically the resource users. Users add tags to online items, such as images, videos, bookmarks and text. These tags are then shared and sometimes refined. A general review of social bookmarking tools, one popular use area of folksonomies, was given in the April edition of *D-Lib* [1]. In the article the authors elaborate on the approach taken by social classification systems and the motivators behind tagging. They write, "...tags are just one kind of metadata and are not a replacement for formal classification systems such as Dublin Core, MODS, etc.... Rather, they are a supplemental means to organise information and order search results."

In this article we look at what makes folksonomies work. We agree with the premise that tags are no replacement for formal systems, but we see this as being the core quality that makes folksonomy tagging so useful. We begin by looking at the issue of "sloppy tags", a problem to which critics of folksonomies are keen to allude, and ask if there are ways the folksonomy community could offset such problems and create systems that are conducive to searching, sorting and classifying. We then go on to question this "tidying up" approach and its underlying assumptions, highlighting issues surrounding removal of low-quality, redundant or nonsense metadata, and the potential risks of tidying too neatly and thereby losing the very openness that has made folksonomies so popular.

2. The Folksonomic Flaw

Probably the major flaw of current folksonomy systems – and the number one gripe for those happier with more formal classification systems – is that the tagging terms used in those systems are imprecise. It is the users of a folksonomy system who add the tags, which means that the tags are often ambiguous, overly personalised and inexact. Many folksonomy sites only allow single-word metadata, resulting in many useless compound terms; the majority of tags are generally believed to be "single-use"; that is, to appear only once in the database of tags. At present there is little or no synonym (different word, same meaning) or homonym (same word, different meaning) control. The system administrators do not impose judgement about the tags chosen by users. Plural and singular forms, conjugated words and compound words may be used, as well as specialised tags and "nonsense" tags designed as unique markers that are shared between a group of friends or co-workers. The result is an uncontrolled and chaotic set of tagging terms that do not support searching as effectively as more controlled vocabularies do.

Some users do not consider this a problem; they may argue that tags are there primarily to help the particular end-user who is submitting them. In addition, Clay Shirky [2] has argued that in folksonomies there are no such things as synonyms, because users employ tags for specific reasons. Therefore every different user-selected word actually has a unique meaning (e.g., cinema and movies). However, as tagging systems become more popular and critics of the system continue to raise their voices, many in the folksonomy community recognise responding to the charge that there is a "folksonomic flaw" would help make clear that there is great additional public value to be had from the "private" metadata submitted. Optimisation of user tag input, to improve their quality for the purposes of later reuse as searchable keywords, would increase the perceived value of the folksonomic tag approach.

3. User-created Tags

So what exactly are tags? A simple definition would be to say that tags are keywords, category names, or metadata. In essence, a tag is simply a freely chosen set of textual keywords. However, because tags are not created by information specialists, they do not at present follow any ubiquitous formal guidelines. This means that items can be categorised with any word that defines a relationship between the online resource and a concept in the user's mind. Any number of words might be chosen, some of which are obvious representations, others making less sense outside the tag author's context.

Two well-known examples of folksonomy systems, to which we will refer extensively in this article, are del.icio.us™ [3] and flickr™ [4], both services owned by Yahoo. Del.icio.us is a tagging system for URLs that integrates with the Firefox browser by means of bookmarklets (JavaScript interface elements) and offers the user the ability to store and retrieve their bookmarks on the del.icio.us website and to identify each bookmarked URL by tagging with appropriate metadata. Flickr is an online photo storage system that allows users to identify their photographs by means of a set of tags. Each site can be browsed or searched for resources that match a given tag.

In order to understand how we can make tags more searchable it is important to understand users and why they submit certain tags. At this time, little is known about the decision-making process behind tag selection, and quantitative data is relatively scarce. One useful approach might be to examine users' motivations when adding tags, see why they decide on particular words, observe how many tags they add and compare how the same items are classified by different users. It might also be helpful to see how feedback affects tag use and how users modify tags in the light of the behaviour of others. However, such studies take time and resources.

One small-scale study [5] carried out by Ulises Ali Mejias of Ideant raises many interesting points, though it fails to find much in the way of concrete information about why certain decisions are made. One of the conclusions Mejias draws from his study is that although the tags used often have a hidden meaning known only to their creator, there are clearly certain tags (repeated tags) that have a social shared meaning alongside the personal meaning. It is these tags that are seen as offering the most benefit; methods are therefore sought to encourage their creation and use. Although this is clearly an area for future research, more is known about the distribution of tags. For example, it is possible to see the top 50 tags added by users of del.icio.us [6].

Many folksonomy sites offer third-party visualisations of the most popular tag choices in common tagging; there are a number of tools available offering a variety of visualisation methods, including tag.alicio.us [7], extisp.icio.us [8] and facetious [9]. Tag.alicio.us is an experimental design from Olivier Richard that operates as a tag filter, retrieving links from del.icio.us according to tag and time constraints (e.g., tags from this hour, today, or this week). Extisp.icio.us displays a random scattering of a given user's tags, sized according to the number of times that the user has reused each tag, and facetious is a reworking of the del.icio.us database, which makes use of faceted classification, grouping tags under headings such as "by place" (Iraq, USA, Australia), "by technology" (blog, wiki, website) and "by attribute" (red, cool, retro).

Power laws and tag distribution

Adam Mathes, well known for his timely paper on folksonomies, has suggested that tag distribution follows a power law scenario [10]. The most used tags are highly visible so likely to be used by other users, (few tags used by many). Then there will be a large number of tags that are used only by a few users, (many tags used by few). And finally there will be a huge number of tags that are used by just one or two users.

Mathes explains, "Examining this sort of distribution of tag use could give a better indication of whether a folksonomy converges on terms and foster consensus, or if as the user based grows the vocabulary grows at a more even rate, and the distribution of terms flattens, perhaps indicating less agreement."

Tag popularity

Before writing this article, we conducted a study of our own, collecting a sample data set to see if we could determine to what extent popular objections to folksonomic tagging are based on fact. We took a random sampling of tags from both del.icio.us and flickr. (The methodology is described in Appendix 1.) By taking a random set of sample tags from flickr and determining the popularity of each tag, we found a distribution similar to that predicted by Mathes. Popularity of tags decreases very rapidly, the resultant curve falling asymptotically towards y=1, in a characteristic shape (see Figure 1).

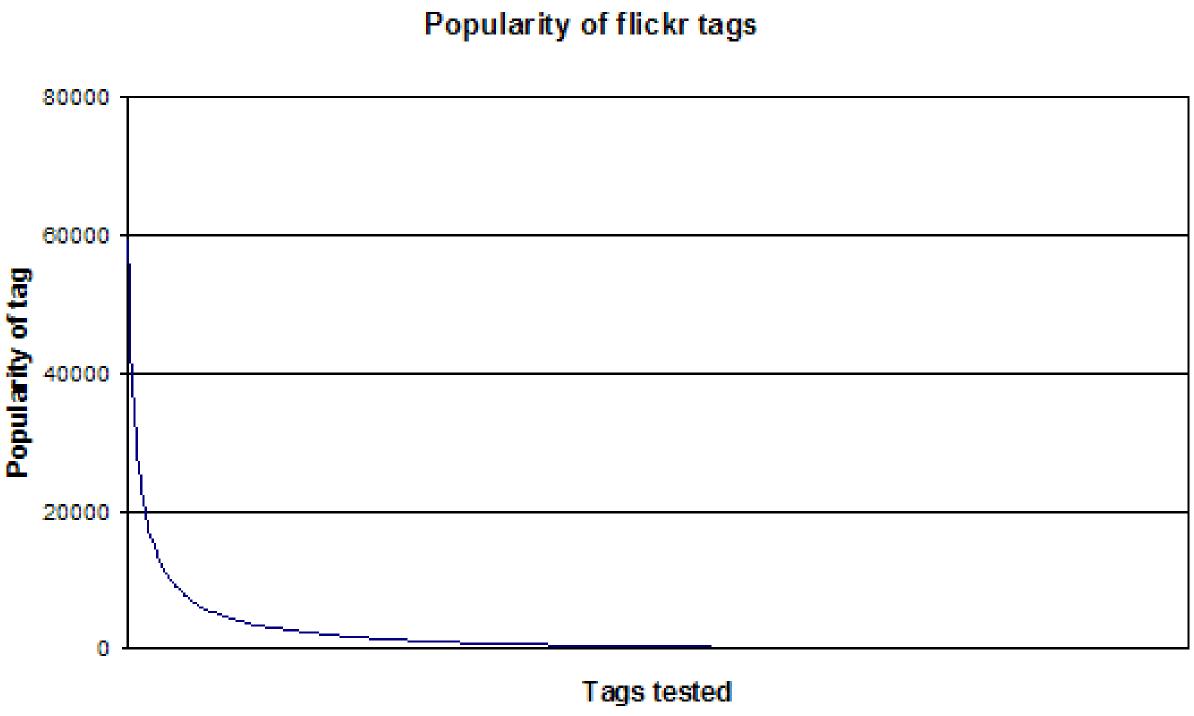

Figure 1: Popularity of randomly sampled flickr tags

Figure 1 demonstrates the extremely large range of tag popularities fairly well. This distribution is often graphed on a logarithmic scale (see Figure 2) in order to compress the larger values onto a reasonable scale and improve readability.

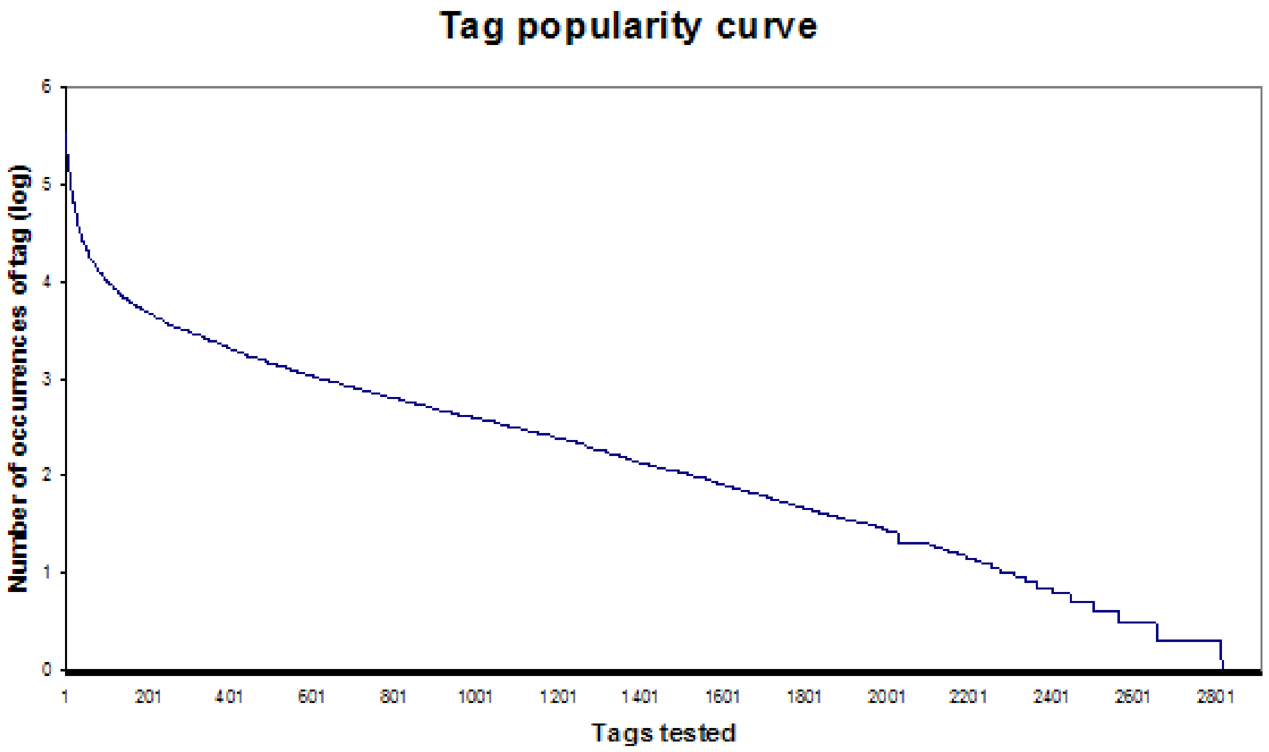

Figure 2: Tag popularity curve

These results indicate that single-use tags do not dominate tagging systems. According to our sample, only ten to fifteen percent of the tags sampled on Flickr and del.icio.us are single-use tags. The expected mass of tags used only once is not represented in these results. This may be due to the possibility that Mathes' text dates from a time when the user population of both sites was somewhat lower. With increasing popularity, the database becomes more deeply littered with unusual tags, misspellings and compound terms, so that the "floor" of the graph may be expected to rise a little with time.

Anecdotal evidence [11] supports the view that there is a natural tendency towards the convergence of tags and that strategies to facilitate this development exist. Stephen Pinker in his text *The Language Instinct* [12] discusses pidgin (a combination of words from other languages absent of any stable grammatical structure) and creole (a combination of words from other languages with a unique grammar imposed) languages. He argues that creole will come from pidgin if people are given the chance to speak to others. It could be argued that similarly social tagging services create the kinds of environments in which we can evolve metadata vocabularies in a natural way.

The evolution mentioned here refers to the production of a single, fairly stable, shared ontology, and this statement is fair in that Pinker's example is limited to a single community. Within a given setting, culture or social grouping, the process progresses as the system reflects currently preferred choices in language, supporting each participant in his or her own contributions to the group.

4. Improving Tag Literacy

Given that there is already a movement towards convergence of tags, how can we foster this trend? At the moment there are two key ways in which the metadata created in folksonomies could be improved to aid searching:

- Educating users to add "better" tags
- Improving the systems to allow "better" tags to be added

Educating users

Currently most users don't give much thought to the way they tag resources, and bad or "sloppy" tags are ten-a-penny in folksonomies. The main casualties are usually enumerated as follows:

- Misspelt tags (e.g., libary, libray)
- Badly encoded tags, such as unlikely compound word groupings (e.g.,TimBernersLee)
- Tags that do not follow convention in issues such as case and number; singular versus plural form (e.g., apple, apples)
- Personal tags that are without meaning to the wider community (e.g., mydog)
- Single-use tags that appear only once in the database. (e.g., billybobsdog)

In order for folksonomies to offer much more in the way of social value, many feel that tag creation needs to becomes a lot more proficient; but are the problems really those described above?

Tagging observed

Returning to the tags we randomly sampled from flickr and del.icio.us, the following flaws could be observed.

- Misspellings, incorrect encodings, and compound words: By testing against multilingual dictionary software, we found that 40% of flickr tags and 28% of del.icio.us tags were either misspelt, from a language not available via the software used, encoded in a manner that was not understood by the dictionary software, or compound words consisting of more than two words or a mixture of languages.
- Words that did not follow system conventions: Almost 8% of the flickr tags and over 11% of the del.icio.us tags were plural forms of words.
- Symbols used in tags: Symbols such as # were used at the beginning of tags, probably for an incidental effect such as forcing the del.icio.us interface to list the tags at the top of an alphabetical listing.

However, we did find that single-use tags were less common than we had expected.

Structures other than dictionary words accounted for a large number of the tags found in our study. Compound words often contained numbers, in constructions such as "17thjuly", or "April11". Conventions have become popular, such as dates represented according to the ISO standard (eg. 20051201 for "1st December, 2005") and the use of the year as a tag. One wildly popular convention is geotagging, a simple method of encoding latitude and longitude within a single tag; this represented over 2% of the total tags sampled on flickr.

A common source of "misspelt" tags was in the transcoding of other alphabets or characters. For example, the umlaut, which is commonly used in German, is usually represented by means of the Latin-1 character set. Since this character set is often unavailable, German users frequently represent an umlaut character by means of a longhand encoding, such as "ue" for "ü". This particular case occurred in several of the sampled del.icio. us tags. Similar technical issues exist with character encoding in several other languages, such as Chinese, Japanese, Russian or Czech. (This leads us to speculate that widespread user adoption of (and increased confidence in) Unicode may be a major factor in the success of folksonomies on the world stage.)

Though the bulk of tags found in our samples from flickr and del.icio.us combined are valid English language dictionary words using US or British spelling, tags from other countries are represented and may be in various foreign languages (see Figure 3).

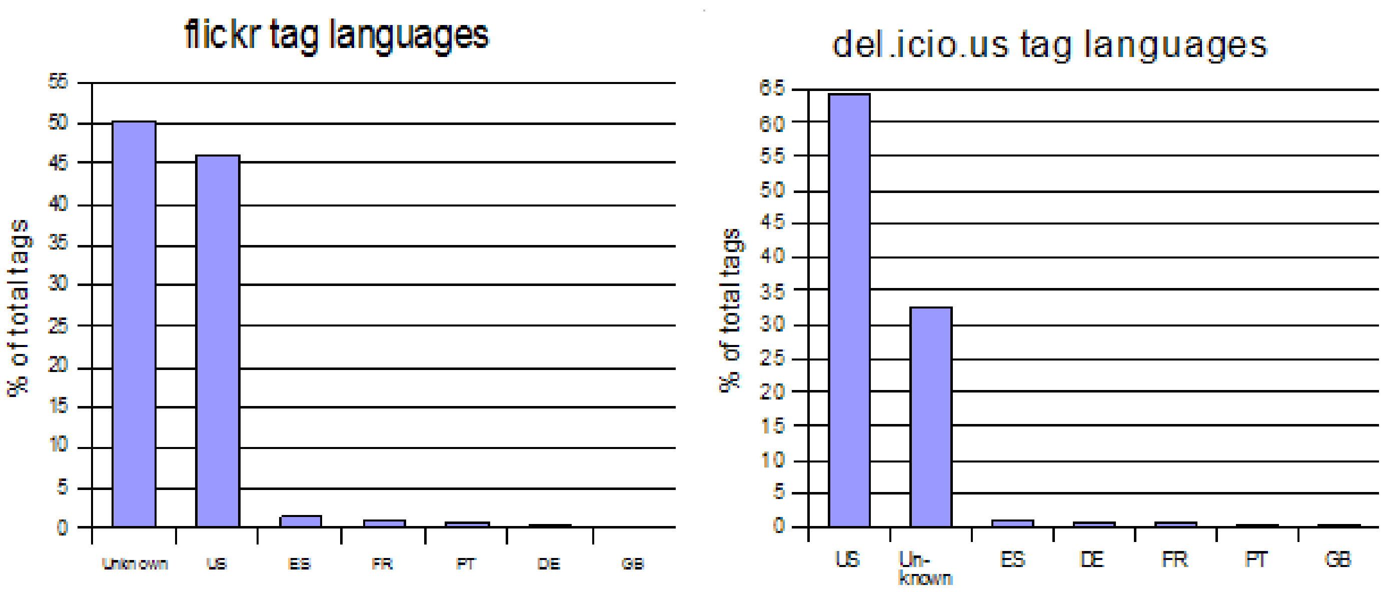

Figure 3: Tag languages estimate from random sample of del.icious and flickr tags

Accurately gauging source language of tags is hampered both by technical issues and by the fact that many words exist in multiple languages, though frequently with differing meaning or grammatical structures; for example the term *son* in English, as in *father-son*, is used in French as a possessive adjective, and in Spanish as a form of the verb *ser*, to be. Fortunately, the majority, almost 90%, of recognised dictionary words used in the samples from both del.icio.us and flickr are nouns.

However, the confusion inherent in folksonomic tagging showed itself most clearly in a feature common to over 10% of all the sample tags taken from del.icio.us – many users attempted to make compound words without simply concatenating words together, but by putting a symbol or a piece of punctuation inside the tag to represent a space. This was particularly interesting, because some users appeared to be attempting to establish a hierarchical structure by building up a "pathway" within the tag. For example, a user tagging several web pages within del.icio.us on the subject of programming languages might tag one topic as "Devel/C++", a second as "Devel/BASIC", a third as "Devel/Perl", and so on.

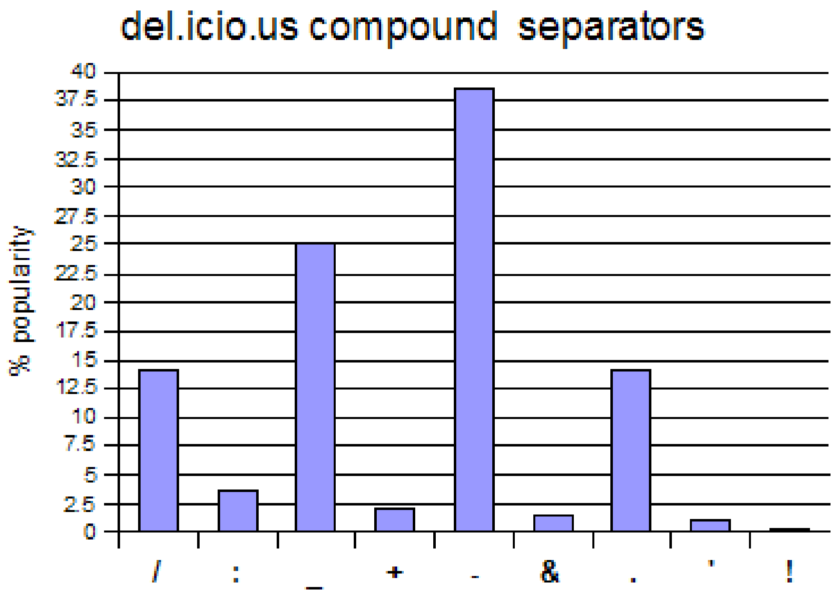

Figure 4: del.icio.us compound word separators

Looking at the variation in symbol chosen for this purpose (see Figure 4), it is clear that no consensus or convention has been chosen by the del.icio.us user community to play the role of the "non-breakable space". Since such compound tags are so common within the del.icio.us tag space, this is one example of how a little guidance might benefit the tagging community as a whole.

Methods for improving tags

As most in the information world will know, improving the quality of user-created metadata is not a new phenomenon. Information specialists have wrestled with the issues involved many times and have suggested various remedies. For example, in an *Ariadne* article entitled "Improving the Quality of Metadata in Eprint Archives" [13], the authors suggest the inclusion of quality assurance processes on creation of the metadata.

To succeed, attempting to improve tag literacy (or tag etiquette) in the folksonomy world involves two processes. Firstly, the community needs to be ready to set rules and agree upon a set of standards for tags. Secondly, users need to be made aware of and agree to follow these rules.

At the moment, although there are no standard guidelines on good tag selection practices, those in the folksonomy community have offered many ideas. Ways in which tags may be improved are presented frequently on blogs and folksonomy discussion sites. In his article on tag literacy, Ulises Ali Mejias suggests a number of tag selection "best practices" [14]. These include:

- using plurals rather than singulars
- using lower case,
- grouping words using an underscore,
- following tag conventions started by others and
- adding synonyms.

Other recommendations from Mejas and others working in this area are that users try to "think specific and general at the same time" and that personal tags are fine as long as more generic tags are also used. The consensus among those in the folksonomy community is that extra tags are always better. Many folksonomies allow users to modify their tags, and there is considerable scope for users to tidy up the entries that they have already created. Although this issue has been widely discussed, any attempt to introduce a "best practice" for users of tagging sites almost certainly requires the enthusiastic participation of site developers and administrators. One might make a case for establishment of a consortium of some of the most popular folksonomy sites, to release a list of general guidelines.

Tagging could be improved by providing users with a set of helpful heuristics that promote good tag selection, such as a checklist of questions that could be applied to the object being tagged, in order to direct the tagger to various salient characteristics. Another idea that could be implemented is to introduce structure within tags. Currently, tags are generally defined as single words or compound words, which means that information can be lost during the tagging process. Single-word tags lose the information that would generally be encoded in the word order of a phrase. This is particularly seen in English, with the dissociation of adjectives from noun. For example when tagging a photo I might want to use tags to describe a black cat and white dog. Once the single-word tags "black," "cat," "white" and "dog" are assimilated into the database, their meaning is lost. Users searching no longer know which animal is black and which is white. However, the problem of adjective/noun dissociation is not equally relevant to every language. In some languages the issue is avoided or mitigated, as in those languages, such as Russian or German, that impose noun and adjective declination for case. With regard to compound words, private conventions are chosen by individuals for indicating relationships within an otherwise flat namespace, but these indications are applied for personal use, are not standard and cannot therefore be leveraged to any common advantage.

The commonness of compound tags, including tags that concatenate more than two words, may suggest that users miss the richness of the sentence structure. If this is indeed the case, a wise solution might include choosing and imposing an authorised method of concatenating terms – the "non-breakable space" mentioned earlier. Although many compound words are produced using separator characters, such as this_is_a_tag or i+tag+therefore+i+am, or by separating words by means of CamelCase (that is, formatting each word with an initial upper-case letter to improve readability), as of today, there are a large number of tags that are literal concatenations of words, which are very difficult to parse usefully, such as "thisisaspecialtag". Splitting these tags currently invokes particular difficulties for developers, analogous to the difficulty posed to machine translation by languages allowing compound words, such as German, Finnish or Greek [15]. Encouraging their replacement would improve the potential for richer search functionality, such as searching for tags including a given noun or adjective, or at least increasing the reliability of guesses as to the language in which the tag itself is written.

Another interesting area for consideration is that of tag bundles. This is the tagging of tags that results in the creation of hierarchical folksonomies. Many have talked about how folksonomies need to evolve through links to more formal systems. As Louis Rosenfeld explains: "In fact, it's exciting to consider how these two approaches might fit together and function as a whole. Neither works especially well on its own: controlled vocabularies often miss out on input from content authors and become rigid, stale, and distant from the vernacular of users; folksonomies will begin to break down for the reasons mentioned above. Treating them as major parts of a single metadata ecology might expose a useful symbiosis: encourage authors and users to generate folksonomies, and use those terms as candidates for inclusion in richer, more current controlled vocabularies that can evolve to best support findability" [16].

Several del.icio.us taggers have established a private pseudo-hierarchy of terms, by establishing tag conventions that resemble directory structures, such as, `Programming/C++`, `Programming/Java`, `Programming/XHTML`. Furthermore, many taggers on del.icio.us have chosen to tag URLs with other URLs, such as the base web address for the server (e.g., a C# programming tutorial might be tagged with `http://www.microsoft.com`). It's difficult to disagree with the idea. When one tags a photograph, one usually includes the location as one or more of the tags. When one tags a digital resource, one might use the organisation that one associates with the tag, which on the Web corresponds to the general location of the tag – and what could be more logical than to refer to an organisation by what one knows to be a unique identification string, their web address? Nonetheless, there is something absurdly recursive about this tagging practice.

Smart systems

Alongside educating users, there is much that system creators can do to improve the end-data their systems are helping to create. There are two main ways in which improvements can be made. Firstly, much can be done at the point at which new resources are contributed to the system. Simple error-checking potentially accounts for a number of tag errors – although rather fewer misspellings occur than may be expected. Furthermore, some sites already make tag suggestions when users submit resources. Scrumptious, a recent Firefox extension, offers popular tags for every url [17]. Systems could easily suggest synonyms, expansion of acronyms, and the like when users type in their tags.

Secondly, improvements can be made in the way systems search for resources already in the system. Synonym suggestions could also be made here, suggesting, for example, "ladybug" instead of "ladybird".

One missed area of opportunity is that of more discussion tools through which users can share reasons for tagging things in a certain way. At the moment there is little discussion on folksonomy sites about the appropriateness of tags. Most of the sites do not offer the opportunity to provide actual text feedback, though some allow you to change other users' metadata. Some systems also provide very little indication as to the characteristics of the tagger whose work is on display; more user profiling might help improve browsing – for example, the tagger's preferred language is a valuable piece of contextual information. More understanding of who is submitting certain tags could possibly alter your own personal rating of posts (e.g., "Bob submitted that, he's into PHP and he seems like a good guy, so I'll assume it's useful"). One site looking at user profiling is Collaborative Rank [18], which ranks people based on how helpful and timely their suggestions are.

There are obvious dangers in establishing a positive feedback loop where potentially unsuitable tags may be reused due to the tag's initial popularity and subsequent exposure as a tag recommendation. This leads one to wonder whether it is preferable to have popular (but perhaps not intuitively obvious) tags, or to have a larger spread of relatively uncommon tags, possibly representing more accurate reflections or a wider spread of points of view. In folksonomies as elsewhere, the probable answer is, "it depends".

5. Putting It All into Place

Now that we have a few ideas for tag improvement, is it time that we considered testing them in practice?

Examining tag use and the eventual convergence or consensus on certain terms is undoubtedly an absorbing exercise. However, implementing real-world strategies based upon these assumptions should be approached cautiously, as there is one significant limitation that must be recognised: the various individuals who supply and use tags are geographically and culturally diverse. The strength of a folksonomic approach is often described to be its openness, the ability of any given user to describe the world as he or she sees it. Could one expect a useful consensus to be reached? Is consensus desirable in a tagging system? In tag-based systems, there are at least two stakeholder groups: those who contribute metadata in the form of tags, and the consumers of that metadata. These may overlap; however, there is no reason to assume that a metadata consumer must be familiar with the metadata submission process. While the contributors' choice of vocabulary may have been "trained" by the various means discussed in this article, the metadata consumer may not have had the benefit of that process.

Clay Shirky notes:

> *Tagging gets better with scale. With a multiplicity of points of view the question isn't "Is everyone tagging any given link 'correctly'", but rather "Is anyone tagging it the way I do?" As long as at least one other person tags something the way you would, you'll find it – using a thesaurus to force everyone's tags into tighter synchrony would actually worsen the noise you'll get with your signal. If there is no shelf, then even imagining that there is one right way to organise things is an error.* [19]

Is it possible that by attempting to tidy up tags we are losing the very hook, attraction, or essence of folksonomies?

Folksonomies are popularly related to the anthropological study of "folk taxonomies", a favoured study of cognitive anthropologists in the 1960s, but the significance of this snippet of information is often eclipsed by today's perception of folksonomies as a popular mechanism for creating user-populated search databases. Briefly revisiting the origins of the term is useful, if only to situate the discussions presented here with respect to their antecedents.

A folk taxonomy is most easily defined by contrast to a scientific taxonomy, a naming system to be applied objectively, independently of social matters. Scientific taxonomies, such as the Linnaean taxonomic system, are to be applied independently of personal feeling on the matter. The emergence of the "folk taxonomy" recognised common names as worthy of mention, serving useful functions within a social and cultural context, and the study of folk taxonomies remained popular for some time. However, few generalisable results were extracted from this work, and the work tended to focus on artificially simplified and often trivial semantic domains [20]. It was eventually re-framed as a stage in the study of knowledge structures, consensus and understanding within groups.

Later work from a number of domains provides some insight into the problem domain, but the field is complex, encompassing culture, language and thought. On some details agreement has been reached; people do appear to think in terms of domains [21], and dialect is an indicator of social class, educational level and age.

The subset of a language used in a certain setting (the situated nature of vocabulary choice and manner of speech) is both fascinating and confounding. In internet terms, this is most commonly encountered in the form of "speech communities", groups of people who share a certain set of vocabulary or jargon.

The strengths and weaknesses of folksonomies within classification systems are emergent from the nature of speech within context. Thomas Hardy's poem *An August Midnight* [22] benefits from his dialect:

> *On this scene enter – winged, horned, and spined –*
> *A longlegs, a moth, and a dumbledore*

He might have instead written "A crane fly, a moth and a bee", had he been willing to foresake the opportunity to instill a little local colour, but his choice to use dialect or common names was inspired, and the poem benefits from it. However, a search engine would not. Unless armed with a synonym dictionary able to relate the longlegs with the Harry Long Legs, the father long-legs, the daddy-long-legs (and ignore either of the spiders which share the name!) and the family Tipulidae, Hardy's richness of vocabulary will almost inevitably result in a little-used tag with low social value – that is to say, a tag with little usefulness as a search term.

Even worse, Hardy's unusual vocabulary has been eclipsed in terms of meaning; without a way of injecting a little context into a search for Hardy's bumble-bee, a way of persuading the system that we are looking for the animate but non-human insect known as "a dumbledore", rather than the Headmaster of Harry Potter's Hogwarts, the bumblebee signature is unlikely to be searchable through the noise created by JK Rowling's creation. Hardy has become the victim of a word collision; you can check by asking Google. Searching for the ambiguous "Dumbledore" provides 1,870,000 hits. Providing a little context, we might search for "Albus Dumbledore" for the gentleman wizard, producing 434,000 results and "moth dumbledore" for the insect, receiving a lowly 758 results, most of which are relevant to our poet.

Interfaces designed to breed out such tags, dialect, uncommon, archaic or conflicting terms, are an attempt to build a stable, robust and clearly defined taxonomy of user-provided terms. An analogy might be drawn between this and the various attempts to reform the English language, such as the advent of *Received Pronunciation English*, or the various fashions for words of Saxon or Norman origin. Although often well-meaning, such reforms have historically proven to be a matter of fashion rather than advancement, and cannot be relied upon to produce a stabler form of the language.

6. Conclusions

The investigations described in this article are brief, simple and relatively unscientific, as are the numbers provided within. That the results from both del.icio.us and flickr tended to be rather similar imply that they can be trusted only as much as a short, seat-of-the-pants, peripheral analysis ought to be. Only those with direct access to the del.icio.us and flickr databases can be aware of the exact state of affairs and how it has changed across the months. Curious readers are advised to perform their own investigations. For our purposes, the interesting features of the tags are not in the precise percentages of usage, but in the choice of tag, the choice of structure, and the choice of language. Somewhere around a third of tags were indeed "malformed", in that they were beyond the grasp of a multilingual spell-checker for one reason or another. Many of these were not misspelt, but mis-constructed, some of the latter in a correctable manner.

Still, possibly the real problem with folksonomies in not their chaotic tags but that they are trying to serve two masters at once; the personal collection, and the collective collection. Is it possible to have the best of both worlds? At the moment, many investigations of tag data are in progress, including how tags can be used for searching. As a consequence, development in this field tends to confine itself to methods for improving the quality of the user-contributed tags for this purpose. In practice, this involves promoting commonly-chosen tags above single-use or infrequently used tags by various means, such as user interface enhancements, synonym use and so on. It is possible that the data collected through folksonomy tagging is more complete than we had imagined. Achieving more from that data may be a question of developing an appropriate set of algorithms; in other words, revisiting the data with another aim in mind might reveal usefulness in some categories of "sloppy" tag. Some single-use tags are explicitly designed as such, such as the latitude/longitude markers used by geotagging (flickr). Some may be perceived as valuable or helpful to the reader. Some may be infinitely helpful for search purposes, if only the information provided therein is accessed in an appropriate manner. Is it therefore preferable, rather than attempting to stamp out single use or sloppy tags, to suggest that each item be tagged with a mixture of approaches, including several search-friendly keywords?

Can we be certain that training the user into a relatively restricted choice of tags is purely beneficial, if it is even possible? It is not unlikely that further uses for folksonomy metadata will arise. In other words, is our myopic focus on tag-based search systems leading us to consider metadata such as single-use tags as useless, as automatically "bad" tags, when these tags may yet turn out to have a use in another domain or context? Do such tags have a value to one or another stakeholder, beyond their use as search terms? These are topics for further experiment and observation.

As the various systems making use of tagging evolve, "sloppy" tags can be weeded out. Interface changes can be made to discourage certain practices, such as the use of symbol prefixes to force a tag to the head of one's tag list, and to encourage others, such as the use of a standard method for those who do wish to create tag phrases. Arbitrarily-formed compound words and misspelt tags may become less frequent, and better handled by the search interface.

But as the community making use of each tagging system grows in size and diversity, other problems arise. Systems intended to suggest common or popular tags are trained to promote the hegemony of tags arising from its earlier user population; to search effectively, new users may be required to guess at conventions no more obvious to them than the formal taxonomies that the folksonomy replaced. Improving usability across cultures necessitates recognition of the issues that language, dialect and jargon represent. Discouraging users may mean that they simply do not bother to tag further resources.

The answer is to remain open minded and look at solutions that retain as much as possible of the metadata submitted, bearing in mind that metadata can be mined in all sorts of ways. Amy Gahran of Contentious observes that "A folksonomy merges, diverges, and evolves much the way language does, through usage and interaction" [23]. This is one of folksonomy's great strengths. There is a real danger that by tidying up tags we are condoning the implementation of a destructive solution that may lose valuable metadata. The two questions we need to ask ourselves may be: Even assuming that such a consensus were possible, do we really want a world where everyone speaks a collaboratively defined analogue to the Queen's English? To what extent, in this instance, with a fantastically complex and valuable database of user contributions from all over the world, is it possible to separate the metaphorical baby from the bathwater?

Appendix 1: Methodology

In this article, we sought to determine to what extent the popular objections to folksonomic tagging are grounded in fact. Therefore, a necessary first step was the collection of a sample data set with which to work. In order to establish this data set, sample tags were taken from del.icio.us and flickr as follows:

A number of usernames were collected from the "Most recent updates" view of each website. These were used to access the RSS feed of each user's tags, where such existed, or the web listing, in the event that it did not. The resulting tags were then collected together. A random subset of around three thousand of these tags were chosen, and the number of instances of each tag was then calculated.

This methodology clearly does not provide information about the popularity of the most frequently used tags, since it is likely that only a tiny percentage of the most frequently used tags are represented in the subset chosen. However, the less popular tags were of more interest for the purposes of this article.

In order to check the spelling of tags, we made use of the common Unix tool, aspell [24], checking each tag against several dictionaries by means of a Perl script. For tags that validated successfully in English, the word class was then determined using Princeton University's Wordnet lexical reference system [25].

The accuracy of these results could have been improved further by making use of a word stemming system, particularly in languages other than English, for example, by making use of the Perl Lingua::Stem module [26].

References

[1] Hammond, T., Hannay, T. Lund, B., Scott, J., (2005) Social Bookmarking Tools A General Review , *D-*

Lib Magazine, April 2005, Volume 11 Number 4. <doi:10.1045/april2005-hammond>.

[2] Ontology is Overrated: Categories, Links, and Tags, <http://www.shirky.com/writings/ontology_overrated.html>.

[3] Del.icio.us™, <http://del.icio.us/>.

[4] Flickr™, <http://www.flickr.com/>.

[5] Ideant: A del.icio.us study - Bookmark, Classify and Share: A mini-ethnography of social practices in a distributed classification community. <http://ideant.typepad.com/ideant/2004/12/a_delicious_stu.html>.

[6] The top 50 tags added by users of delicious, <http://del.icio.us/popular/>.

[7] tag.alicio.us, <http://frenchfragfactory.net/ozh/archives/2004/10/05/tagalicious-a-way-to-integrate-delicious/>.

[8] extisp.icio.us, <http://kevan.org/extispicious>.

[9] facetious, <http://www.siderean.com/delicious/facetious.jsp>.

[10] Folksonomies - Cooperative Classification and Communication Through Shared Metadata - Adam Mathes <http://www.adammathes.com/academic/computer-mediated-communication/folksonomies.html>.

[11] Jon Udell's screencast on del.icio.us <http://weblog.infoworld.com/udell/2005/03/14.html>.

[12] Pinker, S., *The Language Instinct: How the Mind Creates Language*. Harper Perennial Modern Classics; November 1, 2000.

[13] Guy, M., Powell, A., & Day, M. (2004) Improving the Quality of Metadata in Eprint Archives, *Ariadne* Issue 38, <http://www.ariadne.ac.uk/issue38/guy/>.

[14] Ideant: Tag Literacy, <http://ideant.typepad.com/ideant/2005/04/tag_literacy.html>.

[15] Koehn, P. and Knight, K. Empirical methods for compound splitting, *Proceedings of the tenth conference on European chapter of the Association for Computational Linguistics*, April 12-17, 2003, Budapest, Hungary.

[16] Folksonomies? How about Metadata Ecologies? <http://louisrosenfeld.com/home/bloug_archive/000330.html>.

[17] Scrumptious, <https://addons.mozilla.org/extensions/moreinfo.php?id=738>.

[18] Collaborative Rank, <http://collabrank.org>, and Michail, A., Collaborative Rank: Motivating People to Give Helpful and Timely Ranking Suggestions (work in progress), last updated: April 23, 2005. <http://collabrank.web.cse.unsw.edu.au/collabrank.pdf>.

[19] Shirky, C. 2005. Ontology is Overrated; Categories, Links and Tags. <http://www.shirky.com/writings/ontology_overrated.html>.

[20] Keesing, Roger M. 1972. Paradigms Lost: The New Ethnography and the New Linguistics. *Southwestern Journal of Anthropology* 28(4):299-332.

[21] Romney, A., Moore, C., and Brazill, T. 1998. Correspondence analysis as a multidimensional scaling technique for non-frequency similarity matrices. Pages 329-345. In *Visualization of Categorical Data*. Edited by Joerg Blasius and Michael Greenacre. San Diego: Academic Press.

[22] *An August Midnight*, <http://www.poetryconnection.net/poets/Thomas_Hardy/16365>.

[23] Amy Gahran's comments on Technorati tags: Good idea, terrible implementation, <http://www.intuitive.com/blog/technorati_tags_good_idea_terrible_implementation.html>.

[24] GNU Aspell, <http://aspell.sourceforge.net/

[25] WordNet, <http://wordnet.princeton.edu/>.

[26] Lingua::Stem, <http://www.nihongo.org/snowhare/utilities/modules/lingua-stem/>.

(January, 17, 2006, Reference [1] was corrected to include Joanna Scott as the fourth author of the referenced article.)

Top | Contents
Search | Author Index | Title Index | Back Issues
Editorial | First article
Home | E-mail the Editor

doi:10.1045/january2006-guy