Concept Modeling by the Masses: Folksonomy Structure and Interoperability

Csaba Veres

Department of Computer Science, Norwegian University of Science and Technology
Csaba.Veres@idi.ntnu.no

Abstract. The recent popularity of social software in the wake of the much hyped "Web2.0" has resulted in a flurry of activity around folksonomies, the emergent systems of classification that result from making public the individual users' personal classifications in the form of simple free form "tags". Several approaches have emerged in the analysis of these folksonomies including mathematical approaches for clustering and identifying affinities, social theories about cultural factors in tagging, and cognitive theories about their mental underpinnings. In this paper we argue that the most useful analysis is in terms of mental phenomena since naive classification is essentially a cognitive task. We then describe a method for extracting structural properties of free form user tags, based on the linguistic properties of the tags. This reveals some deep insights in the conceptual modeling behavior of naive users. Finally we explore the usefulness of the latent structural properties of free form "tag clouds" for interoperability between folksonomies from different services.

Keywords: Web2.0, folksonomy, interoperability, tagging, concept modeling.

1 Introduction

There is currently a great deal of activity revolving around applications and initiatives on the World Wide Web that fall under the rubric of Web2.0, the live Web, social software, or architecture of participation [1]. While there is a great deal of hype and cynicism concerning the phenomenon [2], there is nevertheless some consensus on an interesting set of properties that loosely define prototypical Web2.0 applications.

One important hallmark of Web2.0 applications is that they tend to be based around web services so that there is no requirement to install a special application on a client machine. This already introduces a new dynamic to the application space since functionality can change incrementally and with an extremely fast life cycle. It is reported that Carl Henderson, leading developer at Flickr[1] (a leading Web2.0 application) commented that "on good days, Flickr releases new versions every half an hour"[2]. This development model is complemented

[1] http://www.flickr.com/

[2] http://blogs.warwick.ac.uk/chrismay/tag/flickr/)

D.W. Embley, A. Olivé, and S. Ram (Eds.): ER 2006, LNCS 4215, pp. 325–338, 2006.

by an architecture in which constant evolution makes sense; the "architecture of participation". Such architecture is exemplified in eBay's services which are entirely dependent on the participation of its members, and increasing levels of participation can enable the gradual implementation of feature refinements. Amazon.com is another service where users add value by default: whenever they purchase multiple items, add reviews, add items to the wish list, and in general simply "use" the system, they contribute data which in aggregate can improve the service to other customers by providing recommendations and associations which would otherwise not exist. The web service is then improved to make use of the accumulating data. The architecture is designed by default to improve the service simply as a side effect of its ordinary use, and the improved service is quickly rolled out through program updates.

This participatory architecture enables the harnessing of collective intelligence by aggregating user data, which is the second hallmark of Web2.0 applications: the primacy of data over application. A hallmark achievement in this vein is WikipediA, a brave experiment in creating a collaborative encyclopedia which, ideally, anyone could contribute to. Amazingly this radical departure from the kind of authoritarian editorial style one might expect for a reference of this sort, proved to deliver a product comparable to the most venerable Encyclopedia Britannica[3]. A similar challenge is being laid to traditional news services by the activity of blogging, and services like digg. These activities aided by search tools like Technorati together with syndication and other tools like RSS and trackbacks, make it possible for news and opinions to be disseminated and discussed very rapidly.

The benefits of collective intelligence gained through social interaction have come into the popular limelight through the introduction of services like del.icio.us[4], Flickr, CiteUlike[5], Yahoo MyWeb 2.0 Beta[6] and Google Base Beta[7], in which content is contributed, aggregated, and categorized through the collective actions of its users. In some cases the content is created by users as with the photographs contributed to Flickr, but mostly they are proxied as in the case of bookmarks on del.icio.us, or scientific references in CiteUlike. In either case extra value is added through the classification and organization efforts of multiple users. All of these services employ some form of user annotation of the resources, usually referred to as *tags* (e.g. in del.icio.us, Flickr), but sometimes called *labels* and *properties* (in Google Base Beta). The primary value of these services is not simply the addition of content but organization of content in a way that allows its discovery. Crucially, the system of classification and discovery is not driven by sophisticated organizational and search strategies, but by a network of associations that emerges in the process of opportunistic user behavior.

---

[3] http://www.nature.com/nature/journal/v438/n7070/full/438900a.html
[4] http://del.icio.us/
[5] http://www.citeulike.org/
[6] http://myweb2.search.yahoo.com/
[7] http://base.google.com/

For example on the social bookmarking service del.icio.us, users mark up their favorite web sites with their chosen tags. The service requires a user account, and acts in the first instance as a web based repository for each individual user's bookmarks for their favorite web sites. The web sites are indexed by URL and described with a textual description which is typically generated from the title in the web site. As a result, most bookmarks to the same URL will have the same descriptive title, but this is not necessarily the case because users are free to insert their own descriptions. In addition, users annotate each bookmark with metadata in the form of any number of single word tags. The user interface provides access to popular tags for a given URL at the time of bookmarking, assuming of course that other users have tagged that URL. In addition, users can view other URLs annotated with a particular tag they might use. Because the aggregated "tag use" of all users is available in various forms on the service, users can derive value from each others behavior. For example popular tags for a given URL can influence a user who is also adding that URL to their bookmarks, because popular tags are, putatively, useful for other users. On the other hand, users can find new web sites by following links that were tagged with the same terms as the current one of interest. As pointed out in [3] the novel feature of services like del.icio.us is not their reliance on keywords in lieu of taxonomies for indexing – that idea has been around for years. Instead, the novelty is the immediacy of the feedback from the community of users: "Feedback is immediate. As soon as you assign a tag to an item, you see the cluster of items carrying the same tag. If that's not what you expected, you're given incentive to change the tag or add another ... you can adapt to the group norm, keep your tag in a bid to influence the group norm, or both." The benefits to indexing are that resources are grouped according to flexible category structures that are not imposed by authority. This emerging categorization activity that results from the combination of a large number of users tagging resources for their own use has been called *folksonomy* (e.g. [4]). The most fundamental unit of analysis of tagging on del.icio.us is the *tag set* that each individual assigns to an individual resource, which gives rise to a *tag cloud,* the combined set of tags all users assign to that resource weighted by frequency. A tag cloud is therefore a multiset in which order is ignored but multiplicity is significant.

Such a complex network of data lends itself to analysis in a number of different forms. One obvious approach is to use any number of mathematical techniques for the analysis of complex networks, or to find clusters in multi dimensional spaces (e.g. [5]; [6]; [7]; [8]).

In the following section we will briefly present some select observations about mathematical properties of "tag space". But we argue that such analyses are not enlightening as an explanation for the way tags are used to classify resources. Instead we argue that a cognitive perspective, which looks at the linguistic behavior of tags, can provide a useful explanatory account of tag use. Our analysis suggests that naive users produce tags which display latent properties that are typical of complex conceptual modeling activities. In section 3 we describe an approach that can uncover the latent structure in sets of tags. In section 4 we show that the explicit representation of this latent structure can facilitate interoperability. Finally we conclude in showing that we have strong evidence for sophisticated concept models in spontaneous, un solicited naive user tags, which reflect fundamental properties of the cognitive apparatus.

2 Some Mathematical Observations

[6] presents a thesis on the (by now well known) observation that the distribution of the relative popularity of tags in tag clouds approximates a power law function. Individual URLs tend to have a few popular tags (usually less than 10 in number) which are consistently used by a vast majority of users. [6] argues that there is a shift in the precise function that is approximated by the tag cloud, since the popularity of particular tags can vary due to cultural factors such as the spread of new terminologies. But, while this is undoubtedly true in some cases, [7] show on the basis of a large empirical sample that the shape of tag clouds tend to be remarkably stable. In analyzing historical trends for the most popular tags used for a given URL by an ever-increasing number of users, they make the following interesting observations:

"One might expect that individuals' varying tag collections and personal preferences, compounded by an ever-increasing number of users, would yield a chaotic pattern of tags. However, it turns out that the combined tags of many users' bookmarks give rise to a stable pattern in which the proportions of each tag are nearly fixed. Empirically, we found that, usually after the first 100 or so bookmarks, each tag's frequency is a nearly fixed proportion of the total frequency of all tags used." ([7], p. 6).

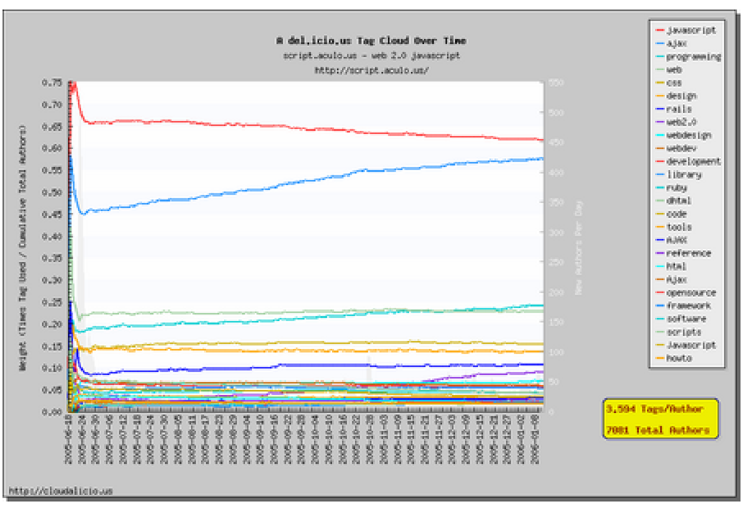

Fig. 1. Tag cloud of del.icio.us tags for the web site script.aculo.us

An example showing this stability, as well as the approximate power law curve, and some evidence for cultural influence in terms of the community uptake of the term "Ajax" is shown in figure 1. ("Ajax" is represented by the slowly ascending, second curve from the top.)

There are several possible explanations for the manifest stability, including relatively un-interesting ones concerning the user interface to del.icio.us, which suggests existing popular tags to each user who tags for their own use a site which was already bookmarked by others. But [7] make an additional observation which shows that such explanations cannot completely account for the observed stability in tag use since the less popular tags which are not shown as suggestions through the interface, display the same stability over time. They conclude that "Shared knowledge among taggers may also account for their making the same choices." Thus while the most popular tags in figure 1 evolved in full view, as it were, of the taggers, the mass of less popular tags at the bottom of the graphic evolved in private. But the two sorts of tags are indistinguishable in terms of their pattern of use, suggesting that the "shared knowledge" contributes significantly to tagging behavior. In addition, [9] suggests an interesting thought experiment. "Suppose I am a really rich guy who wants to influence tags on del.icio.us. So I pay 10000 people to tag resources according to my schema. I tell them to mark one site with 'eek', another one with 'woo hoo', a third one with 'grumpy grumpy head', and so on. With enough people, these should become the most popular tags. But how long will the dominance of these tags last? This is an experiment that does not really need doing!". The simple point, of course, is that the user interface suggestions are popular tags which somehow reflect the shared knowledge discussed in [7].

In this paper we try to find the nature of that "shared knowledge". To do so, we adopt the position that folksonomies are an abundant source of free, interesting data which can give a clue about the way humans organize knowledge, and about the extent to which the mentalistic organizational systems are shared. We subscribe to the hopefully non-controversial position that mental architecture fundamentally shapes our perceptions and organization of the world in which we live. Perhaps more controversially we argue that essential aspects of the mental architecture are fixed and therefore shared by all humans (e.g. [10]) The empirical questions then become "What are the characteristics of the shared architecture?" and "To what degree are they shared?". Clearly there are points of difference in individual conceptualizations. I say 'Library of Congress', but Clay Shirky wants to say 'LOC' [11]. Good for him. But pity the poor soul who calls it 'the square root of a banana'! The point is that the mind creates categories, because that is what minds do. These categories allow some degree of variation, but differences are tightly bounded. The mental architecture enforces the range of possible ontologies and taxonomies that we can bring to bear on the understanding of our universe. All humans share fundamental aspects of mental architecture and therefore properties of possible taxonomies and folksonomies. Folksonomy, on this view, becomes an invaluable source of data for studying the mental processes of naive human classifiers. Conversely, properties of the mental architecture as known from independent sources should give us insight into communal tagging behavior.

A second source of evidence that communal tagging is constrained by, and therefore displays properties of, deep cognitive processes comes from looking at so called *narrow folksonomies* which are to be contrasted with the *broad folksonomies* we have been considering up until now. Tim Vanderwal coined the two terms to describe the two styles of tagging that can be observed on different web applications [4]. The typical behavior on del.icio.us is that many users tag each resource, whereas on the photo sharing service Flickr the default behavior is that tags can only be added by the original contributor and their invited contacts. As a result narrow folksonomies do not display the rich collection of tags that we saw with broad folksonomies. But if we are correct in our claim that the emergent stability of tag clouds with broad folksonomies is due to cognitive facts more than to social, cultural or user interface issues, then there ought to be similar constraints on the tags observed in narrow folksonomies. Some support for this is found through the clustering feature offered by Flickr, which identifies groups of pictures which tend to be associated with overlapping tags, probably using k-means clustering methods. For example a search for clusters with the word "love" returns several distinct groups with tag groups such as {heart, red, valentine, valentinesday, nature, pink, flowers, hearts, white}, {couple, kiss, wedding, bw, people, friends, bride, groom, romance, marriage}, {dog, cat, cute, smile, happy, pet, puppy, cats, kitty, kitten}, and {family, mother, baby, child, kids, fun, daughter, christmas, children, mom}. Clearly the clusters are meant to identify overall themes in the picture collections that can be used to organize photographs. The observation that such clustering is possible suggests that people tend to tag pictures for personal use with sufficient consistency to allow aggregation in a useful way, even though each individual is tagging from their own point of view in complete ignorance of the other users. On the other hand, while clustering is a popular way to process tags for enhanced usability, they have an inherent limitation in that they conflate many dimensions simultaneously [12]. The cognitive approach will give us a way to keep these dimensions distinct.

3 A Cognitive Approach

The hypothesis that folksonomies contain hidden properties that are also observable in formal taxonomies was investigated by [13]. Inspired by the cognitive theory of Lawrence Barsalou[8] [14] and the linguistic insights of Anna Wierzbicka [15], he described a distinction between purely taxonomic concepts and a number of other categories of concepts which were not taxonomic. The idea is that taxonomic concepts are those which describe the basic entities in the world, and

[8] We realize that there is a vast literature on human categorization that we are not covering here, as kindly pointed out by an anonymous reviewer. We leave these out mainly because the theoretical underpinning as elaborated in [13] would gain little by their inclusion.

can be represented in the customary generalization hierarchies where each level in the hierarchy contains disjunctive categories whose members resemble one another more closely than they do members of other categories at the same level. Further, members of a category on a given level are also members of a category at all higher levels. Perhaps most importantly, membership in a category allows a large number of inferences to be drawn about entities. A clear example is in the domain of animals: cats resemble each other more than they resemble dogs, and all cats (and dogs) are also mammals. Cats can be further specialized as Siamese cats and Russian Blue, where Siamese cats resemble one another more than they resemble Russian Blues, and so on. There are also a very large number of inferences that can be drawn about an individual if it is known to be a cat. We know its rough dimensions, its weight, appearance, that it needs food, goes to the toilet, likes to breathe air, and so on.

On the other hand there are a large number of categories which do not display these properties. Consider as an example the class which is described by the word *weapon*. If someone tells you that their country just acquired a fantastic new weapon, what can you conclude about the acquired object? For sure, it can be used to inflict harm and destruction. But how big is it? Is it solid or gas? Is it even a substance, or is it instead a kind of psychological weapon? Does it look like a pistol? Or an inter continental ballistic missile? Or a dog? In point of fact, very little can be inferred from category membership, except its functional property. Wierzbicka [15] calls these concepts purely functional ones, because they describe heterogeneous types which can be used to fulfill a particular function. In addition, she describes three other non-taxonomic categories as follows.

A second kind of category, exemplified by *furniture*, is formed because its members are often experienced together in a *common location and serving a common function*. Furniture can refer to a very loose and heterogeneous collection of "things" which might include tables, chairs, lamps, ashtrays, stereo systems, televisions, and any number of other items with very little resemblance to one another. A third kind of category that also depends on exemplars being *collected in a common location* is exemplified by *groceries* and *dishes* (as in "go wash the dishes"). In addition to being united by a common location, exemplars of these categories share a *common explanation for their collective existence*, or a *common origin: groceries* can include anything put in a shopping basket at the supermarket including non food items, and *dishes* can refer to any food eating implement used for a meal including plates, pots, knives and forks. This latter example is also interesting because it shows that ambiguity of the word *dishes*: in its taxonomic use it can refer only to different kinds of dishes used for serving food, such as *cereal bowl*, *salad bowl*, and so on; but in its collective use it can also refer to pots and pans and forks. It is possible for elements of this sort of category to lose their collective status as long as they retain a temporal bond. For example *leftovers* can be scattered in various locations but the concept still retains its collective status by virtue of the fact that there was some time and place for their common place of origin. Finally, there is a category whose exemplars have *similar sources and similar purposes or functions*, but aren't necessarily experienced together in a common collection. This sort of category includes *vegetables*, *medicines*, and *herbs*. For example *vegetable* describes a heterogeneous collection of entities that people grow in the ground to be used for food. Members of this category acquire an unusual interpretation when used in plural form: "I had three vegetables for dinner" would seem misleading if I had three carrots, whereas "I had three birds for dinner" would be fine if I had three quails.

One important feature of these types of categories, as we have already hinted, is that they can be distinguished on their grammatical properties. This is not only theoretically interesting but also practically useful because it makes possible the automatic discovery of the appropriate type of category that a given term represents. A comprehensive set of grammatical tests for distinguishing the categories is detailed in [13]. As an example of distinguishing between two different types by their grammatical properties, consider the following sets of sentence frames. *Functional* category names display the following pattern of acceptable and unacceptable (*) frames:

– a toy/vehicle/weapon
– toys/vehicles/weapons, three toys/vehicles/weapons, many toys/vehicles/weapons
– * a lot of toy/vehicle/weapon
– a lot of toys/vehicles/weapons
– * much toy/vehicle/weapon

whereas *functional collocations* exhibit the following pattern of frames (note they are almost, but not completely identical to mass nouns in their pattern of use):

– * a furniture/cutlery/clothing
– * furnitures/cutlerys/clothings, *three furnitures/cutlerys/clothings, *many furnitures/cutlerys/clothings
– a lot of furniture/cutlery/clothing
– *a lot of furnitures/cutleries/clothings
– * much furniture/cutlery/clothing
– an item of furniture/cutlery/clothing

[13] used these categories to compare the structure of the semi formal taxonomies used in YAHOO directory and DMOZ to categorize a resource, with the set of tags assigned by users to the same resource. He found surprising similarities, indicating a similar distribution of the category structures in tags and in the directories. But an interesting difference was a disproportionately large use of taxonomic concepts in the user tags. This is sensible if we assume that the directory categories exist mainly to collect heterogeneous unknown resources according to various function related criteria. In contrast, taxonomic classifications are about single types, so the taxonomic classifiers are likely to be used more often as tags where the resource is already known and a specific view can be taken about their type. In the directories which are used for resource discovery it makes sense to commit to this sort of classification less frequently. In fact in the rare circumstance that taxonomic categories are used, they tend to be leaf nodes where the narrow categories are more appropriate.

4 Folksonomy Interoperability

The cognitive approach provides a way that latent structural information can be extracted from user tags in a given service. But if the cognitive processes are ubiquitous, then their impact should be observed in all applications that utilize user tags. We should therefore be able to achieve interoperability of tags across different applications. Tom Gruber, the author of possibly the most often cited definition of Ontology, considers two possible scenarios from a future Web2.0 where this would be beneficial; first, users might wish to interoperate different services on which they have independently tagged content, and second, search engines might be able to exploit user tags on different services to produce better search results[16]. In both scenarios the key is interoperability of tags such that no one application has precedence over another in terms of tag reference. If tags from different sources are to be compared in some way, then there must be an explicit agreement on the interpretation of the possible patterns of tag use. To solve this problem Gruber suggests an ontology of tags in which the representation of each tagging instance requires at least a four place relation: Tagging(Object1, tag1, tagger1, source1). We could then have n-tuples of the form

- Tagging(Object1, tag1, tagger1, source1)
- Tagging(Object1, tag2, tagger1, source1)
- Tagging(Object1, tag1, tagger2, source1)
- Tagging(Object1, tag3, tagger3, source2)
- Tagging(Object2, tag1, tagger4, source2)

on which a set of axioms can be defined. These axioms might address questions like tag equivalence, for example. So, tagger1 might tag Object1 as both tag1="san francisco" and tag2="sanfrancisco". Are tag1 and tag2 identical? There is obviously not an absolute right answer to this, but an explicit assumption could be stated in terms of axioms defined in the ontology. Then one could go on to ask, if tag1=tag2, does this mean tagger1 only assigned one tag to Object1? Once again assumptions made by implementations can be explicitly stated in axioms. This proposal is about establishing the relationships between individual tags, which precedes the more interesting possibilities for tag based resource discovery. While tackling the issues of syntactic equivalence, synonymy and ambiguity of tags is clearly important, the question of interoperability more broadly construed should include notions of semantic similarity.

Suppose as a concrete example that you had several web services in regular use, each annotated by a set of tags, and you wanted them to inter operate. For example you could be writing a document on Writely[9] the web based word processor, which allows users to annotate documents with tags, and you wanted to collect a set of relevant URLs from del.icio.us and a set of relevant photographs from Flickr. Suppose the document was about the impending bird flu epidemic in 2006, and you wanted relevant links and photos for the different content areas in the paper. Searching for *bird flu* on the two services gives the results in

[9] www.writely.com

figure 2. The del.icio.us tags are obtained from the list of "common tags" that are returned from a search for "bird flu". The clusters that are returned from a similar search on Flickr are shown on the left side of figure 2.

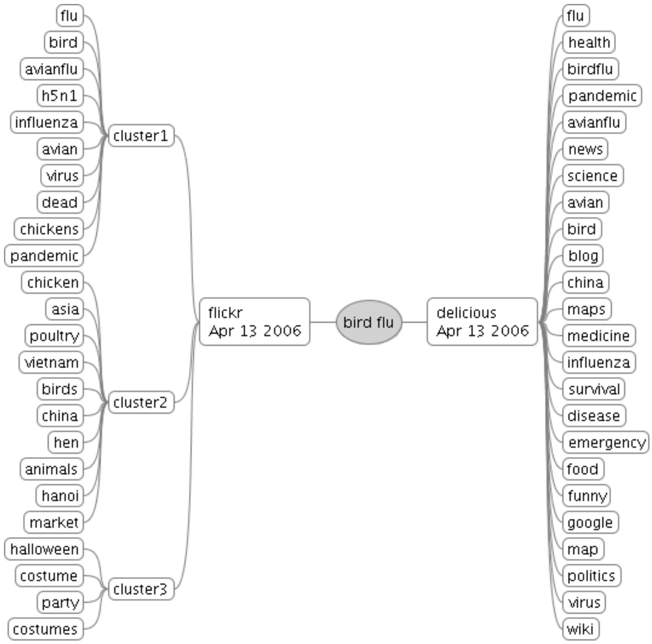

Fig. 2. Tags returned from Flickr and del.icio.us from a search on “bird flu”

It is quite apparent from looking at the set of returned tags that a wide variety of topics are related to *bird flu* in both services, and retrieving all of the resources on the basis of the single tag will give too many irrelevant results for a particular interest. The problem is in identifying the relevant resources in each service and to match them across services. The large number of tags in figure 2, mixed in terms of focus and generality, makes it difficult to find meaningful connections even for humans. The situation is obviously worse for automated processing. For example, cluster2 on Flickr appears to be more about travel, or geography, than the bird flu. How is an automated process to make sensible connections? Which cluster will contain photographs to match the content on del.icio.us?

Our claim in this paper is that tags on both services contain a latent structure which explains their cognitive associations to the various resources, and provides some semantics for the associations. Exposing this structure will clarify the ways in which various tags relate to one another across applications. The process for exposing the latent structure involves a number of steps of natural language processing, and the details are beyond the scope of this paper. (A forthcoming paper will detail this process). However, a brief summary is given here. First, tags are categorized according to a rough division according to the primary grammatical categories Noun, Proper Noun, Verb, Adjective/Adverb. This requires a number of heuristics to resolve ambiguities when they arise. Then the Nouns are further sub divided according to the categories outlined above. Currently this involves manual grammaticality decisions using the templates discussed earlier, but work is well underway toward automating the process. The outcome of the process is the division of the tags into a number of distinct grammatical/semantic categories which are shown with a human interpretable label in figure 3. For example the grammatical category of nouns which describe entities united by a common function are labeled as a category of "related things with common uses, roles". Similarly, taxonomic categories are labeled "What kind of thing is it?".

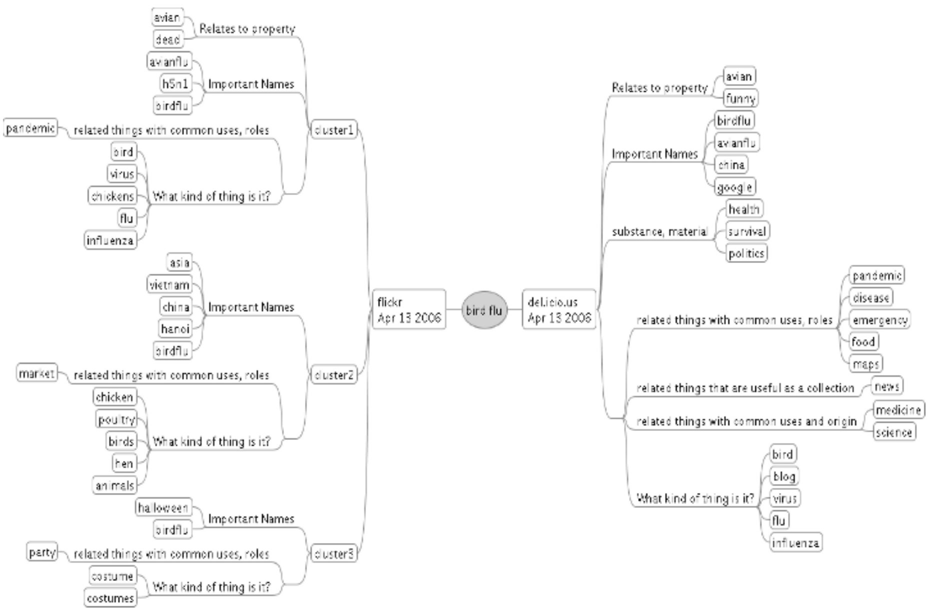

Fig. 3. Tags categorized by grammatical/semantic type

The grammatical/semantic categories tend to group the tags in sensible groups which makes their relationship to the resource, and each other, clear. For example the taxonomic tags on del.icio.us {bird, blog, virus, flu, influenza} refer to specific kinds of entities that are involved in the function/role connected events {pandemic, disease, emergency, food, maps}, and have come to be referenced with the common names {birdflu, avianflu, china, google}. The categories group tags in semantically distinguished relations to the resource. The inclusion of terms like "google" in the original tag set is odd, probably reflecting the perception that Google somehow plays an important part in our awareness of the disease. But our categories help with this, telling us that Google is not one of the things that the tag *bird flu* is about; it is about *birds*, *viruses*, and *diseases*.

Let us now compare tags across the two services by trying to match the tags in the available categories.

1. Taxonomic, which tells us specific kinds of things. Four out of five elements in the del.icio.us set match tags in cluster 1 of Flickr, only one matches cluster 2, and none match cluster 3. If we try to expand the matching process by supplementing each tag with synonyms and more/less general terms from a resource like WordNet, we discover that all of the tags in Flickr cluster 2, and *bird* in del.icio.us fall in the same hierarchy of terms. It is important to note that using WordNet in the comparison process is simplified because the linguistic categories reduce the number of terms that need to be compared. We can therefore specify more precise search phrases on both resources, based on the two matching groups of taxonomic terms, which will retrieve two sets of matching results. {virus + flu + influenza} and {bird + chicken + poultry + hen + animals}
2. Function, which tells us about the uses and roles of the resources. Again if we compare across the two services we find two matches: an exact match on *pandemic*, and a match that can easily be derived through the synonym set for market in WordNet: [grocery store, grocery, food market, market]. The third cluster on Flickr is once again without a match. The search terms again are expanded by inserting into the cluster that was matching in the taxonomic classifications. The sets become: {virus + flu + influenza + pandemic} and {bird + chicken + poultry + hen + animals + market}
3. Names. Again there is a straightforward match between delicious and the two clusters. In addition, there is now a match for cluster 3, so we begin a new set: {virus + flu + influenza + pandemic + birdflu + avianflu} and {bird + chicken + poultry + hen + animals + market + china + birdflu}, {bidflu}

At the end of the matching process we have three sets of tags that identify matching content on the two services. This can be used in several ways, but we

China closes all Beijing poultry markets (Business Week)
China Closes All Beijing Poultry Markets - Yahoo! News.

PandemicFlu.gov
CDC – Influenza (Flu) | Avian Flu
Avian Flu – What we need to know

Fig. 4. Flickr photographs and matching del.icio.us URLs for the search phrases [poultry market china bidflu] and [flu pandemic birdflu]

illustrate with simple search. Submitting all search terms as a conjunctive search yields no results because the phrase is overly specific. We therefore submitted only the most general term from each category of nouns, yielding a small number of resources in both services. Figure 4 illustrates a sample pairing of Flickr photographs and del.icio.us URLs.

The ontology helps inter operation in two ways. First, it reduces the number of nodes to be compared by introducing independent dimensions of comparison. While some correspondences between tags could be established without the linguistic categories, the search space would be much higher, as noted in point 1 above. But equally importantly the semantics of the groupings is uncovered. So, for example, if we have the need to manipulate the search terms as above, we have semantically distinct groupings that can be treated differently. We know for example that proper names do not have more general terms, so we can't exclude any of them. On the other hand we could exclude all taxonomic terms and only use functional ones, to get all markets in China, not just poultry ones.

5 Conclusion

We have argued that folksonomies which reportedly have no structure or constraint on their properties in fact do have rich structure, determined by the properties of our mental faculties. We have shown a method that can expose significant aspects of that structure, together with a semantics that can be used to construct an ontology from the folksonomy basis. This ontology, we argue, is a simple conceptual domain model built through an unconsciously mediated collaboration. Finally, we showed a way in which the ontologies can facilitate interoperability between application dependent tag sets.

The work described in this paper has lofty goals, but is described in the spirit of the emerging Web2.0. That is, the content that is needed for the complex operations is collected as a default behavior of system use. Value can be added to existing applications without first solving all the complex problems. Enhancing manual operation of tag based services is only first goal. Once the data is available, it can be used to research more complex problems addressing automation and, eventually, bootstrapped into enabling interoperability in the most complex Semantic Web applications.

Acknowledgments. This work was sponsored by the Norwegian Research Council, WISEMOD project, 160126V30 in the IKT-2010 program.

References

1. Levy, S. and Stone, B. The New Wisdom of the Web. *Newsweek,* April 3, 2006. http://www.msnbc.msn.com/id/12015774/site/newsweek/
2. Boutin, P. Web 2.0 The new Internet "boom" doesn't live up to its name. Slate. Posted Wednesday, March 29, 2006, http://www.slate.com/id/2138951/
3. Udell, Jon. Collaborative knowledge gardening. InfoWorld. August 20, (2004). http://www.infoworld.com/article/04/08/20/34OPstrategic_1.html

4. Vanderwall, T. Explaining and Showing Broad and Narrow Folksonomies, http://www.vanderwal.net/random/entrysel.php?blog=1635, February 21, (2005)
5. Shen, K. and Wu, L. Folksonomy as a Complex Network. Computer Science, abstract cs.IR/0509072. http://arxiv.org/abs/cs.IR/0509072, (2005)
6. Speroni, P. On Tag Clouds, Metric, Tag Sets and Power Laws http://blog.pietrosperoni.it/2005/05/25/tag-clouds-metric/ (2005)
7. Golder, S., and Huberman, B. A. The Structure of Collaborative Tagging Systems, http://www.citebase.org/cgi-bin/citations?id=oai:arXiv.org:cs/0508082 (2005)
8. Grigory Begelman, G. Keller, P. and Smadja, F. Automated Tag Clustering: Improving search and exploration in the tag space. Collaborative Web Tagging Workshop, 15 th International World Wide Web Conference, Edinburgh, Scotland, 2006.
9. Veres, C. Emerging Patterns. http://csabaveres.net/blog8/?p=7 February, 7, (2006)
10. Jackendoff, R. Semantics and Cognition Cambridge, Mass. MIT Press (1983)
11. Shirky, C. Matt Locke on folksonomies. March 01, 2005. http://many.corante.com/archives/2005/03/01/matt_locke_on_folksonomies.php
12. Hearst, M. Clustering versus Faceted Categories for Information Exploration, Communications of the ACM, 49 (4), April 2006
13. Veres, C. The Language of Folksonomies: What tags reveal about user classification. In Natural Language Processing and Information Systems. Proceedings of the 11th International Conference on Applications of Natural Language to Information Systems May 31 - June 2, Klagenfurt, Austria. Springer, LNCS 3999. (2006)
14. Barsalou, Lawrence W. Deriving categories to achieve goals. in Bower, G. (Ed.) The Psychology of Learning and Motivation: Advances in Research and Theory, Academic Press, 1991.
15. Wierzbicka, A Apples are not a 'kind of fruit': the semantics of human categorization. American Ethnologist 313–328 (1984)
16. Gruber, T. Ontology of Folksonomy: A Mash-up of Apples and Oranges. http://tomgruber.org/writing/ontology-of-folksonomy.htm#_edn4, Jan 19, (2006)