Semantic Annotation from Social Data

Geir Solskinnsbakk and Jon Atle Gulla

Department of Computer and Information Science
Norwegian University of Science and Technology
Trondheim, Norway
{geirsols, jag}@idi.ntnu.no

Abstract. Folksonomies can be viewed as large sources of informal semantics. Folksonomy tags can be interpreted as concepts that can be extracted from the social data and used as a basis for creating semantic structures. In the folksonomy the connection between these concepts and the tagged resources are explicit. However, to effectively use the extracted conceptual structures it is important to be able to find connections between the concepts and not only the already tagged documents, but also new documents that have not previously been seen. Thus, we present in this paper an automatic approach for annotating documents with concepts extracted from social data. This is based on representing each tag's semantics with a *tag signature*. The tag signature is then used to generate the annotations of documents. We present an evaluation of the approach which shows promising results towards automatic annotation of textual documents.

1 Introduction

The last years we have seen a growing amount of social services on the web. Amongst these are a wide range of collaborative services that offer users the possibility of tagging a multitude of resources. These resources can be anything on the web, ranging from images, videos to documents. These services can aid the user in organizing information by letting the user attach tags to the resources for easy access at a later time. In addition, the social aspect lets users share resources and tags, so that others can also take advantage of the effort each individual user puts into tagging. There exist many tagging systems, like Flickr (http://www.flickr.com) which lets users share and tag images, Delicious (http://www.delicious.com) which lets users tag and share any resource specified with a URL, Bibsonomy (http://www.bibsonomy.org) which lets users tag and share literature references. Users are free to choose which tags to apply to resources with no centralized control of the vocabulary. The networked data structure resulting from such systems are often referred to as Folksonomies [1].

Tags in folksonomies can be seen as a basis for concept extraction for semantic data structures, which can also be seen in several publications lately [2, 3]. The conceptual structures are one side of the story, however, it is also an interesting problem to connect the concepts (tags) with documents on the web.

This is especially interesting for applications that require search and browsing of the structure and documents. On one hand, we already have a mass of manual annotators (the users of the folksonomy) who generate annotations. Unfortunately, the users have not tagged every single document. This means that there is a huge amount of documents that have not yet been annotated by folksonomy users. Although the documents have not been tagged by users, the documents may be interesting for a browsing facility. Determining the correct annotation of a document automatically is thus the problem we are targeting in this paper. As a solution towards this problem we propose an approach towards fully automatic annotation of documents that have never been seen by the system (i.e. documents that have not yet been tagged by any user). Since we are working on folksonomy data we will use the terms tag and tagging rather than concept and annotation, respectively, for the remainder of the paper. Tags on their own carry only limited semantics. However, we can exploit that the folksonomy can be seen as a large repository of informal semantics to extend the semantics of the tags. This is done by associating each tag with a *tag signature*. The signature takes the form of a vector of semantically related terms, which are weighted to describe the strength of the relations between the tag and the terms in its vector. The tag signature is constructed based on the (textual) resources that have previously been tagged by the users of the folksonomy. By utilizing the tag signatures for suggesting tags to documents, we are using the content (or topic) of the document and tag signature to suggest tags. Thus our approach is not only able to suggest tags to resources that have been tagged before, but also to resources which are new to the system. The approach is evaluated (using training and test data) based on a data set crawled from Delicious. The results of the evaluation are promising in terms of automatically assigning tags to documents.

The remainder of the paper is organized as follows: Section 2 gives an overview of the related work, while Section 3 gives an overview of tag signatures and the approach for automatic tag suggestion. Section 4 describes the evaluation and results, followed by a discussion of our findings in Section 5. Finally the paper is concluded in Section 6.

2 Related Work

The related work for this paper is directed at tag recommender systems, since these systems essentially provide some of the same functionality that we are targeting.

Mishne [4] presents an approach for suggesting tags for weblog posts. This is done by first finding similar weblog posts using information retrieval techniques. The tags used on the most similar posts are retrieved and ranked before being presented to the user. Another system for tagging of blog posts is described by Qu et al. [5]. The system uses key phrase extraction applied to the blog content to find tags which can be applied to the blog post. The system described by Baruzzo et al. [6] also uses key phrase extraction for generating tag recommendations to the user. The keyphrases are extracted from the text and mapped

to domain ontology. Spreading activation is employed in the ontology to locate common ancestors which are presented to the user as new tag recommendations. In [7], Lipczak et al. present an approach based on a combination of extracting candidate tags from the resource and using information found in the folksonomy. Candidate tags are found from the title and the URL of the resource, tags related to the resource, and tags related to the user.

Musto et al. [8] apply a combination of content-based and collaborative-based approaches to generate tag recommendations. The content-based approach analyzes the resource to tag, and extracts candidate tags from the URL, the HTML title and meta tags. The candidates are scored by taking into account type of source (URL, title etc.) and the occurrence frequency within each source type. The collaborative approach searches an underlying corpus of users, resources, and tags to find candidate tags. Finally the user is presented with tags from one or both of the candidate tag sets based on some strategy. Jäschke et al. [9] present two different algorithms for tag recommendation based on folksonomy data. The first is based on collaborative filtering, and the second is based on the FolkRank algorithm. Gemmell et al. [10] describe an approach for tag recommendation based on adapting the k-nearest neighbor algorithm to folksonomy data.

Most current methods use either the content of the resource (key phrase extraction), or the data found in the folksonomy as a source of tags to recommend to the user. Our approach to automatic tagging is based on a combination (even though we do not extract tags from the content). We use the information in the folksonomy (the mapping from tag to resource) and the content of the resource to build a semantic representation of each tag. In this way our approach is able to suggest tags (that are used in the folksonomy) to documents that have not been seen before. Systems that purely use the graph structure of the folksonomy to recommend tags, will suffer when trying to recommend tags to a resource not previously seen. On the other hand, systems that purely rely on extracting tags from the content may lead to an increase in the tag vocabulary. Hence, reusing tags that already exist in the folksonomy will ensure that the vocabulary in the folksonomy is consolidated.

3 Tag Signatures

Users that contribute within a community to tag and share resources on the web generate what is often referred to as a folksonomy [1]. Folksonomies consist mainly of three entities; (1) users; (2) tags; and (3) resources. Bookmarking is the action of a user attaching one or more tags to a specific resource, and the combined data is called a bookmark.

Heymann views this data as *triples*[11] {user, tag, URL}. The interpretation of the triple is that *user* has applied *tag* to the resource identified by *URL*. As the user has actively engaged in applying the tag(s) to the resource we make a basic assumption that the tag(s) make up a description of the documents' content. In terms of the user, the tag(s) applied signal the semantics of the resource and

should be representative for the resource's content, so it is later easy to find (both for the user himself and others in the community).

The assumption made above is used as a basis for generating an extended semantic representation of the tags using the contents of documents to which a tag has been applied. This representation associates each tag in the folksonomy with a vector of semantically related terms. Each term is given a weight that reflects the importance of the term with respect to the tag. This means that a term can be connected to several different tags, but with different weighting, signaling that the term has a different importance with respect to each tag. We refer to our semantic representation as a *Tag Signature*. Two different considerations are made when deciding how to weigh the terms in each tag signature. The first is that the weight should reflect the internal semantics of the tag. This means that we want to give a high weight to terms that are good at characterizing important aspects of the tag. The second is that we want the weight to reflect the external semantics of the tag. This in essence means that we want the term to be good for discriminating this tag from others. Thus we apply the $tf \cdot idf$ [12] measure for weighting the terms in the signatures. The collection of terms and their weights collectively represent the semantic content of the tag, and we thus refer to the tag signature as an extended semantic representation of the tag, which greatly extends the pure syntactic representation of the tag. The tag signature materializes as a vector. The definition is given in [13] (in [13] we use the term Tag Vector), but we repeat it here for convenience as Definition 1.Details of the construction of the tag signature can be found in [13].

Definition 1. *Tag Signature. Let V be the set of n terms (vocabulary) in the collection of tagged resources. $t_i \in V$ denotes term i in the set of terms. The tag signature for tag j is defined as the vector $T_j = [w_1, w_2, \ldots, w_n]$ where each w_i denotes the semantic relatedness weight for each term t_i with respect to tag j.*

3.1 Unsupervised Tagging Approach

Unsupervised tagging can be used in many application areas such as tag recommendation, automatic tagging of a set of documents, document classification, etc. Our approach to automatic tagging takes as input an untagged document and returns a ranked list of tags. The similarity between the document content and the tag is based on the tag signature. Since the tag signature is represented as a vector of weighted terms, and similarly the document can be viewed as a vector of weighted terms, we propose to use the cosine measure to calculate the similarity between the two. The calculation is shown as Equation 1 [12], where $w_{i,d}$ is the weight of term t_i in the document, $w_{i,j}$ is the weight of term t_i in T_j, and n is the number of terms. In our implementation, we have stored all tag signatures in a tag signature index, and use the document as a large query into the tag signature index. The list of tags returned can be cut off at top m tags, or at a threshold for the similarity score.

$$sim(d, T_j) = \frac{\sum_{i=1}^{n} w_{i,d} \times w_{i,j}}{\sqrt{\sum_{i=1}^{n} w_{i,d}^2 \times w_{i,j}^2}} \tag{1}$$

Our approach does not increase the tag vocabulary (as for instance key word extraction techniques might do by proposing new tags). This is a benefit since the document will be tagged according to the already used tags. This means that we can classify the documents according to tags that are already used and are found in the semantic structure. However, if the coverage of the tags is not sufficient, it may be the case that new tags have to be introduced. In such cases the system could have as fallback strategy to implement one of the content based tag suggestion algorithms found in the literature. Another benefit is that the extended semantic representation of the tags allows us to adapt the semantics of a tag to the way it has been used by the users. This implies that a tag may have a different tag signature in different communities, since the tags may be used in slightly different contexts. However, this also means that there will be domain restrictions to the approach. For automatic tagging of good quality we are reliant upon a good coverage of the domain.

4 Evaluation

The experiment is performed on a data set from Delicious that we crawled between December 2009 and January 2010. We only kept bookmarks pointing at resources under "http://en.wikipedia.org/wiki/", the English section of Wikipedia. The crawl resulted in 228536 bookmarks created by 51296 users, 72420 unique tags, and 65922 unique URLs. We kept only English Wikipedia documents so that we could map the documents to a dump of Wikipedia (from June 2008) which has been cleaned and Part of speech (POS) tagged [14]. We performed some simple filtering of the crawled data, removing bookmarks pointing at certain document classes. All bookmarks pointing at documents prefixed with *category:*, *user:*, *image:*, etc. were removed from the delicious data set. This filtered 14162 bookmarks. We were able to map the URLs in 91.2% of the remaining bookmarks to the Wikipedia dump, leaving us with a total of 195471 bookmarks. Mapping failures may have been due to encoding problems, articles that have moved, etc. Next, we filtered the bookmarks based on tags. This was done by lowercasing tags and removing all tags that had not been used by at least 5 users and in 25 bookmarks. This is to ensure that we remove some of the noisy tags found in folksonomies, and assure that the tags have been sufficiently used. The final tag set consisted of 2988 tags (used to tag 59610 documents).

The data set has been randomly split into two parts based on the documents, one for generating the *tag signatures* (training set) and one for the evaluation (test set). The training set consists of 29845 documents while the test set consists of 29765 documents. The *tag signatures* have been constructed according to the description given in Section 3. Further we have performed the evaluation using both standard preprocessing and by extracting terms based on POS tags in the Wikipedia collection. The POS based pre-processing is based on extracting only noun phrases from the text, splitting phrases and stemming individual terms.

The first part of our evaluation is designed to find how well the tag assignments made by our approach corresponds with the tags assigned to the docu-

ments by the folksonomy users. This is done by constructing the tag signatures based on the training set and comparing the tag assignments generated by our approach in the test set with the original tag assignments in the bookmarks of the test set. As a simple base line, we have chosen to use keyword search (named KW Tags). The keyword search is performed by using each tag in the folksonomy (same tag set as we use for tag signatures) as a keyword query matched against the document and generates for each document a ranked list of tags which we compare our method to. All indexing and search has been implemented using Lucene[1].

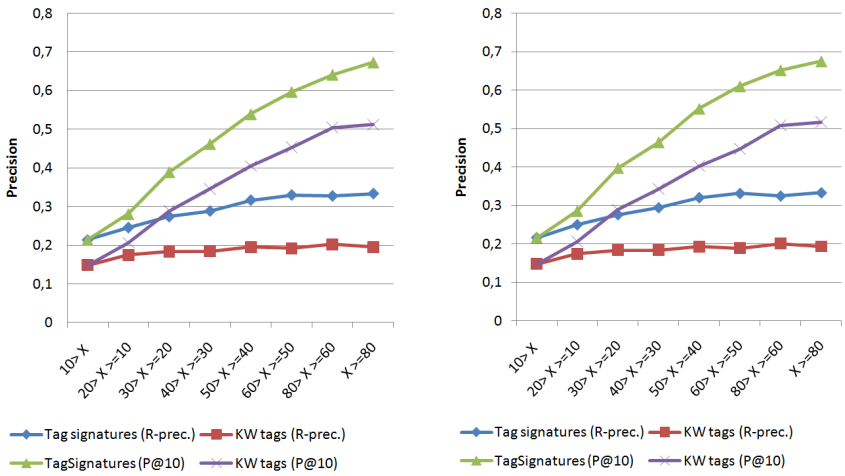
The second part of the evaluation, the user evaluation, has been performed by presenting a group of 6 persons (including one of the authors) with 15 randomly selected documents. For each document the user has been presented with the top 10 ranked KW Tags, and the top 10 Tag Signature based tags (in random order). Tags that have been used to tag each document in the original folksonomy data set have been removed from the evaluation set. Thus we are evaluating only new tag assignments. This is done to learn more about the quality of the tags that are suggested but that have not previously been used to describe the documents. In case of overlap between the two result sets, the list of tags has been padded with extra tags so that the user always is presented with 20 tags. The evaluators used a 5 point scale in which 1 meant that the tag was not appropriate to describe whole or parts of the document content, while 5 meant that the tag was highly descriptive of whole or parts of the document.

4.1 Results

In the first part of the evaluation, we investigate how well our results compare to the tag assignments made by users in the folksonomy. We have used the training set to generate the tag signatures and the test set for evaluation. This means that the text of the documents we evaluate is not incorporated in the training phase. Consequently, the set of bookmarks has been split in two, one for the training set and one for the test set. We have calculated two different measures, the R-precision, and Precision @ 10 (P@10). R-precision for the tag assignments of a single document is calculated by taking all tags assigned to the document by the users (of the folksonomy; the original tag assignments) in the test set as the relevant set of tags, R, with $|R|$ elements. Next we take the top $|R|$ results from KW Tags and our method and calculate the precision in these sets. We also check the precision in the top 10 tags (ranked by the cosine measure) as these tags are the most interesting to suggest to users. We have grouped the results according to the number of unique tags assigned to the documents (Figure 1), the number of times a user has tagged the document (Figure 2(a)), and the size of the documents after preprocessing (Figure 2(b)).

The average R-precision value calculated over all documents in the test set is 0.224 for our approach and 0.155 for the keyword based approach. The average P@10 is 0.238 for our approach and 0.168 for the keyword based approach.

[1] http://lucene.apache.org



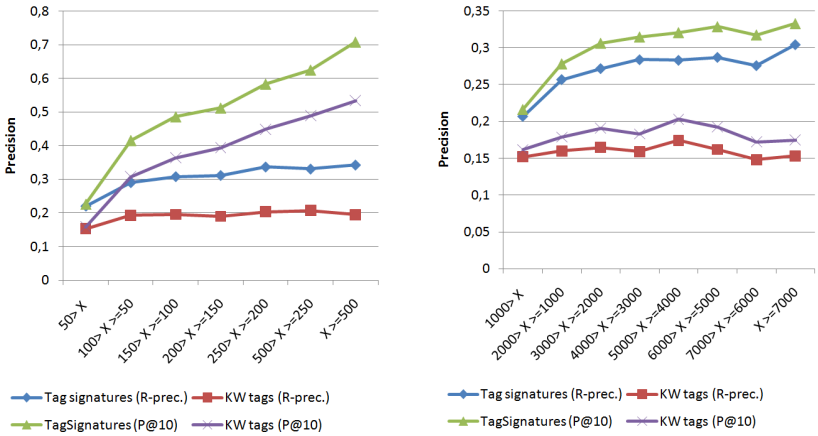
(a) Standard preprocessing (b) POS based preprocessing

Fig. 1. Results grouped by number of unique tags (X) assigned to each document.

Figure 1(a) and 1(b) show the results of KW Tags and our method based on standard preprocessing and POS tag based preprocessing, respectively. The results show that the quality of the two approaches seem quite comparable, thus using the POS information does not improve the quality of the results significantly. Next we can note from the figure that our results are consistently significantly better through all groups than using the pure keyword based approach. Manual examination of the results also shows that our approach is able to find tags that are not present in the document text.

In Figure 2(a) we have grouped the results according to the number of tag assignments to each document. These results show the same trends as the previous graph, as should be expected, since there is a correlation between the number of unique tags assigned to a document and the total number of tag assignments to a document. As the number of tags assigned by users to a document increases, so does the probability of being able to suggest one of these tags. The increase in the experiment metrics with increasing number of unique tags/tag assignments should thus cater for at least parts of this effect.

Figure 2(b) shows the results grouped by document size (after preprocessing). From the graph we can see that our approach scores consistently higher than KW Tags for both measures. From the figure we can see that the results from the KW Tags seems quite stable with only small changes as the size of the documents increase. The approach based on the tag signature on the other hand seems to increase, but with a lower rate as the document size increases as in a logarithmic function. This is an quite interesting result. Since the tag signatures have the form of a vector, we should expect that the number of tags that mach a given document increase as the document size increases (the number of



(a) Number of user tag assignments (b) Document size in number of terms

Fig. 2. The figures 2(a), and 2(b) show the results grouped by the number of user tag assignments and the document size in number of terms, respectively.

potential keywords to match increase). This should also be visible for the KW tags case. However, the results do not show this kind of effect, rather a decrease in the evaluation metric as the number of document terms passes 5000. Thus we interpret this as a result pointing towards that the added semantics of the vectors are able to generate better suggestions.

Figure 3 shows the results from the user evaluation. The data series named *Tag Signatures* is based on the top 10 tags suggested by our approach, while the data series *KW Tags* is based on the top 10 tags suggested by using the existing tags in the system as keyword queries into the documents. Tags that have been used to tag these documents in the folksonomy data set have not been evaluated. Thus the tags evaluated are "new" to each of these documents. The evaluation is performed to check the quality of the remaining tags from the first part of the evaluation, i.e. tag assignments from our system that are not present in the form of bookmarks in the data set collected. The graphs show that the quality of the tags were assessed by the evaluators to be, on average, of higher quality for the *Tag Signature* data series in 10 out of 15 documents. The average value was found to be 3.18 for tags suggested by our approach and 2.91 for tag suggested by the keyword based approach. Although not statistically significant results, we see this as a positive tendency. Manual examination of the documents and tag evaluations showed that there was some disagreement (as can also be seen from Table 1 which shows the standard deviation of the user evaluation scores). This seems to point towards that it is hard to understand the mechanisms that lie behind tagging. It seems that one tag may be valuable to one user, while it is not that valuable to others. The users' intention when tagging (or evaluating a tag in our case) seems to be very important. Some users would like to tag based on the general topic of the document, while others may want to tag based

on certain details in the document. This makes it hard to evaluate tagging on single documents, and our approach seems to be more appropriate when we take a large sample of documents into consideration. Two types of tags our approach seems to not handle satisfactory are subjective tags and very general tags (like interesting, history, etc.). Subjective tags are hard to handle in general and will be discussed further in the next section. Very general or broad terms may cover a very wide topic (like history, which can be used to tag documents about World War II and music history in the 60's). This can however be viewed as a variation on tag ambiguity which we address in the next section.

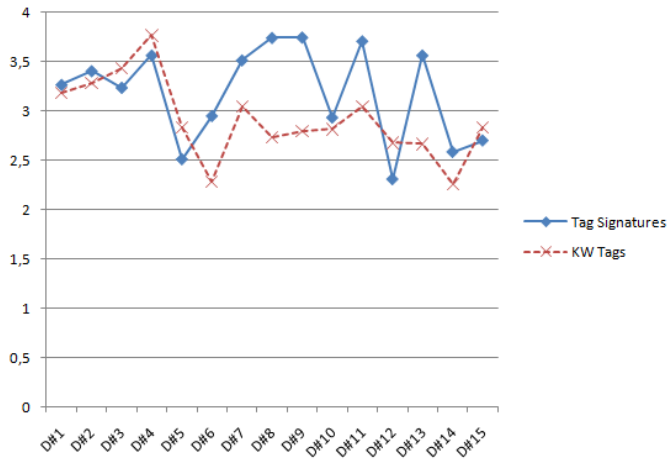


Fig. 3. The results from the user evaluation. Based on 15 randomly chosen documents.

Table 1. The standard deviation of the results from the user evaluation.

Exp./Doc.	D#1	D#2	D#3	D#4	D#5	D#6	D#7	D#8	D#9	D#10	D#11	D#12	D#13	D#14	D#15
$\sigma_{TagSign.}$	1.550	1.379	1.546	1.544	1.601	1.502	1.334	1.198	1.385	1.469	1.160	1.380	1.395	1.544	1.476
σ_{KWTags}	1.525	1.427	1.358	1.280	1.703	1.379	1.455	1.388	1.527	1.266	1.443	1.479	1.481	1.469	1.510

Table 2 shows the tag assignments given by our approach and by the keyword based approach for the document "Comparison of layout engines (HTML5)" (based on the 2008 Wikipedia dump). The results show that the two approaches have an overlap of two tags, *firefox* and *xhtml*. If we polarize tag suggestions as being either good or bad and define good tag suggestions as those with an average score above 3, we see that in our approach 9 out of 10 suggested tags qualify, while for the keyword based approach only 5 out of 10 qualify.

Table 2. Example set of tags for the document "Comparison of layout engines (HTML5)" (2008 version).

Tag signatures		KW Tags	
Tag	**Score**	**Tag**	**Score**
ie	4.5	firefox	4.2
firefox	4.2	xhtml	4.2
xhtml	4.2	engine	3.3
compare	3.7	emulation	2
mozilla	3.8	values	1.8
xforms	3.8	xml	4
webstandards	4.2	input	2
png	1.8	property	1.7
css	4	experimental	1.2
xslt	3.3	internet	3.7

5 Discussion

The results described in the previous section show that our approach using tag signatures for automatic assignment of tags to documents previously not seen by the system has quite good performance. However, when looking at P@10 (average 0.238), we see that we are not able to find all tags applied to the documents by users of the folksonomy. What about the quality of the remaining tags suggested? The second part of the evaluation was supposed to give us an answer to this question, but due to disagreement among the users, it is hard to give a conclusive answer. In fact, the disagreement among the evaluators highlights the problem of evaluating tag assignments. The intention of a user is highly relevant as discussed in the previous section. We found that on average, the score given to tags suggested by our system (3.18) seems to indicate that tags suggested by our system have some positive aspects. Thus although we do not have any conclusive evidence, P@10 would most likely be higher since our system suggests tags that, even though not applied to the document by users in the folksonomy, seem to make sense among the evaluators. For a definite answer to the question of the overall quality of the tag suggestion, we would need to perform a larger evaluation.

One of the strengths of our approach is, in our opinion, that it is able to assign tags to documents that have not been seen by the system previously. We are thus not as bound as methods that adhere to strictly using approaches based on collaborative filtering. The tags suggested are based on the content of documents previously tagged with each tag, and the terms are weighted based on balancing the internal and external representation of the tag. Thus we might say that our approach is a combination of content based and folksonomy based tagging. Further the positive results we have achieved, tell us that the quality of the tag signatures seems reasonable, they are able to describe the characteristics of the tags in terms of a weighted vector of terms.

Tag disambiguation is a concern that we have not addressed in the current phase of our research. Tags have a tendency to be ambiguous (polysemy, homonymy etc.), which is also described in the literature (e.g. in Heymann et al. [15]). Take for instance the tag *apple*. Apple can be used in the computer company sense or in the fruit sense. In our case, tag ambiguity may cause the tag signatures to be imprecise, meaning that they span two or more specific topics (causing drift of the signature). This may have affected our results negatively, by suggesting inappropriate tags to documents. Tag ambiguity can however be reduced by applying one of several measures found in the literature for tag disambiguation (e.g. in Garcia-Silva et al. [16] or Angeletou et al. [17]). In our approach tag disambiguation could be applied during tag signature construction, and would generate several tag signatures (one for each sense) for ambiguous tags. Subjective tags would also give rise to some degree of ambiguity. How do you quantify what *cool* or *interesting* means? These types of tags are hard to deal with in automatic systems, since what one person finds interesting may be uninteresting to another. Thus these types of tags are rather useless to apply in automatic systems, and these kinds of systems should focus on the topic of the document.

6 Conclusions

In this paper we have presented an approach for automatically annotating documents with folksonomy tags using tag signatures. The signatures are materialized as a vector of weighted terms, in which the weights reflect the semantic relatedness of the term with respect to the tag. Our evaluation shows that our approach beats naive tagging, using a direct match between tag and document. We found that we are able to annotate documents in which the tag is not present using the tag signature as a semantic connection. Further, the annotations are not made purely based on what a document has been tagged with in the folksonomy, but takes into account the content of the document as well. The evaluation is based on presenting annotations to documents that have not been seen by the system before and interpret this as evidence that our tag signatures carry more semantics that the tag on its own.

Acknowledgment. This research was carried out as part of the IS_A project, project no. 176755, funded by the Norwegian Research Council under the VERDIKT program.

References

1. Thomas Vander Wal. Folksonomy coinage and definition. http://vanderwal.net/folksonomy.html, Accessed February 8. 2011.
2. Peter Mika. Ontologies are us: A unified model of social networks and semantics. In *The Semantic Web - ISWC 2005*, volume 3729 of *Lecture Notes in Computer Science*, pages 522–536. Springer Berlin / Heidelberg, 2005.

3. Dominik Benz, Andreas Hotho, and Gerd Stumme. Semantics made by you and me: Self-emerging ontologies can capture the diversity of shared knowledge. In *Proceedings of the 2nd Web Science Conference (WebSci10)*, Raleigh, NC, USA, 2010.
4. Gilad Mishne. Autotag: A collaborative approach to automated tag assignment for weblog posts. In *Proceedings of 15th International Confernece on World Wide Web (WWW)*, pages 953–954. ACM Press, 2006.
5. Lizhen Qu, Christof Müller, and Iryna Gurevych. Using tag semantic network for keyphrase extraction in blogs. In *Proceedings of 17th Conference on Information and Knowledge Management*, pages 1381–1382. ACM, 2008.
6. Andrea Baruzzo, Antonina Dattolo, Nirmal Pudota, and Carlo Tasso. Recommending new tags using domain-ontologies. In *WI-IAT '09 Proceedings of the 2009 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology*, pages 409–412. IEEE, 2009.
7. Marek Lipczak, Yeming Hu, Yael Kollet, and Evangelos Milios. Tag sources for recommendation in collaborative tagging systems. In *Proceedings of the ECML/PKDD 2009 Discovery Challenge Workshop*, pages 157–172, 2009.
8. Cataldo Musto, Fedelucio Narducci, Pasquale Lops, and Marco de Gemmis. Combining collaborative and content-based techniques for tag recommendation. In *Proceedings of 11th International Conference on E-Commerce and Web Technologies (EC-Web)*, volume 61 of *LNBIP*, pages 13–23. Springer, 2010.
9. Robert Jäschke, Leandro Marinho, Andreas Hotho, Lars Schmidt-Thieme, and Gerd Stumme. Tag recommendations in folksonomies. In *Proceedings of the 11th European Conference on Principles and Practice of Knowledge Discovery in Databases (PKDD)*, volume 4702 of *LNAI*, pages 506–514. Springer, 2007.
10. Jonathan Gemmell, Thomas Schimoler, Maryam Ramezani, and Bamshad Mobasher. Adapting K-Nearest Neighbor for Tag Recommendation in Folksonomies. In *Proceedings of the 7th Workshop on Intelligent Techniques for Web Personalization and Recommender Systems*, 2009.
11. P. Heymann, G. Koutrika, and H. Garcia-Molina. Can social bookmarking improve web search? In *First ACM International Conference on Web Search and Data Mining (WSDM'08)*.
12. R. Baeza-Yates and B. Ribeiro-Neto. *Modern Information Retrieval*. ACM Press, New York, 1999.
13. Geir Solskinnsbakk and Jon Gulla. A hybrid approach to constructing tag hierarchies. In *On the Move to Meaningful Internet Systems, OTM 2010*, volume 6427 of *Lecture Notes in Computer Science*, pages 975–982. Springer, 2010.
14. J. Artiles and S. Sekine. *Tagged and Cleaned Wikipedia*. Available from http://nlp.cs.nyu.edu/wikipedia-data/, Accessed December 2009.
15. Paul Heymann, Daniel Ramage, and Hector Garcia-Molina. Social tag prediction. In *Proceedings of the 31st annual international ACM SIGIR conference on Research and development in information retrieval*. ACM, 2008.
16. A. Garcia-Silva, M. Szomszor, H. Alani, and O. Corcho. Preliminary results in tag disambiguation using dbpedia. In *CKCaR'09: Proceedings of the 1st International Workshop on Collective Knowledge Capturing and Representation at K-CAP 2009*, 2009.
17. Sofia Angeletou, Marta Sabou, and Enrico Motta. Semantically enriching folksonomies with flor. In *1st International Workshop on Collective Semantics: Collective Intelligence & the Semantic Web (CISWeb 2008) at ESWC*, 2008.