Extracting Time and Location Concepts Related to Tags

Yukino BABA[1], Fuyuki ISHIKAWA[2], and Shinichi HONIDEN[1,2]

[1] The University of Tokyo
[2] National Institute of Informatics

Abstract. Folksonomy is a method of classifying content, and it is widely used in some web services. It allows users to choose tags (keywords or terms assigned to specific content) freely and to search content by referring to the tags. Compared to existing classification methods, folksonomy reflects the users' intention more directly because of its unlimited vocabulary and multiple tags for one content item. Moreover, it has a useful characteristic where tags represent the description of the content. Although tags are intended to be a rich semantic description of web content, machines cannot understand what the tags mean because they are just keywords. We describe a method to extract the concept related to the tag in a machine-understandable way by focusing on the features of content annotated with each tag. In particular, we target the problem of extracting the temporal and spatial concepts of the tags on Flickr, a popular photo sharing service by looking at the date and location distributions of photos for each tag. We evaluated the concept extracting method on a snapshot of actual Flickr data and show that it can identify a tags' concept in a manner similar to the way a person can.

1 Introduction

Folksonomy [1] is a method of classifying content, and it is widely used in web services (e.g., del.icio.us, Flickr, YouTube). It allows users to choose tags (keywords or terms assigned to a content) freely and to search content by referring to tags. Compared to existing classification methods, folksonomy reflects the users' intention more directly because of the unlimited vocabulary and multiple tags for one content item. Moreover, it has a useful characteristic where tags represent the description of content.

Meanwhile, the objective has been to build a Semantic Web where all web content contains machine-understandable metadata describing the meaning of each kind of content. However, the workload required for manual metadata creation and annotation is serious. Tagging could be part of the solution to this problem: A tag is utilized as a description of the content, but the tag is just a textual label, so machines are not able to understand the meaning. For example, when content is annotated with the tag Christmas, machines have no perception of what the tag means.

To solve this problem, we describe a method to extract the concept related to the tag in a machine-understandable way by focusing on the features of content annotated with each tag. For example, from the feature "Photos taken during December 24th and 25th are frequently annotated with the tag Christmas," we can get the information "The Christmas tag is related to a period during December 24th and 25th." In folksonomy, we can make a direct correlation between a tag and the features of content annotated with the tag because tags are added to content directly. Furthermore, we consider that the concepts obtained from tags better correspond to human recognition because of a better reflection of the people's intention, and we are likely to be able to extract the tags' concepts, which are difficult to determine using the top-down definition approach.

Using our method, for example, when content is annotated with Christmas and no date information is added, it is explicit that the content is related to a period during December 24th and 25th. This knowledge is useful to search content not only by keywords but also by time.

In particular, we target tags on Flickr [2], a popular photo sharing service that supports user-generated tags, and we extract tags with time and location concepts from the tags and information on the time and location at which the photos were taken. We define the time and location tags as being those related to time and location (e.g., time tags: August, cherry blossoms, Sapporo Snow Festival, location tags: Tokyo, Daibutsu, Sapporo Snow Festival[3]). Furthermore, we define the temporal and spatial concepts of time and location tags as being the ranges related to the tags (e.g., 2007-02-05 12:21:58 to 2007-02-16 23:48:24, (lat: 35.616279, lng: 139.650307) to (lat: 35.772702, lng: 139.859047)).

Our approach targets not only the tags that are explicitly related to the time or location (e.g., August, Tokyo) but also the tags that are implicitly related to the time or location (e.g., Cherry blossoms, Daibutsu). The latter tags' concepts are difficult to determine using the top-down approach, while the approach considering the features of tagged content as the users' intention is likely to extract the tags' concepts more easily.

We extract the temporal and spatial concepts by analyzing the distributions of annotated times and locations where the photos were taken. A method of determining whether the tags on Flickr are related to time and/or location has already been proposed [3]. However, this research did not extract the concepts of the tags. Our method extracts the concepts of the time and location tags.

The contributions are as follows:

- We provide an approach for extracting the concepts of time and location tags from the feature of the content annotated by the tags.
- We extend the existing method of determining whether the tags are related to time and/or location.
- We describe an application and analysis of this method for actual tags and photo information data taken from the snapshots posted on Flickr.

[3] Each tag can be both a time tag and location tag. Such tags can be considered as "event" tags, but we do not make exceptions for these tags in this paper.

We formalize our problem in Section 2. We describe the existing methods to determine time and location tags in Section 3 and describe a new method to extract the concepts of time and location tags in Section 4. Section 5 discusses the experiment, in which we evaluated our way of extracting concepts. We describe related work in Section 6 and conclude the paper in Section 7.

2 Problem Definition

In Flickr, each photo p has information on the date it was taken t_p and location it was taken l_p and some tags are annotated. We denote a tag generally as x and define certain classes according to the distribution of their temporal and spatial usage patterns as follows:

Time Tag The tag, the temporal distribution T_x of which has more than one dense region.

Location Tag The tag, the spatial distribution L_x of which has more than one dense region.

Moreover, we define concepts of these tags as follows:

The concept of the time tag The dense region in the tag's temporal distribution T_x.

The concept of the location tag The dense region in the tag's spatial distribution L_x.

The procedure of concept extraction is as follows:

1. Determine whether a tag is a time or a location tag; i.e., determine whether the tag has more than one dense region in the time or location usage distribution.
2. Extract the concept of this; i.e., the concept is taken to be the dense region if the tag's usage pattern has more than one dense region.

A standard pattern detection method, Naïve scan [4], perform steps 1 and 2. However, Scale-structure Identification [3], a method of determining whether the tags on Flickr are related to time and/or location, can perform step 1 with higher accuracy. Hence, in this paper, we present a way to apply Scale-structure Identification to step 2. In rest of this paper, we describe two methods to determine 1, Naïve Scan and Scale-structure Identification in Section 3. In Section 4, we show the methods to specify 2 with Naïve Scan and present the method to use Scale-structure Identification for 2, region specification.

3 Determining time and location tags

In this section, we describe two methods to determine whether a tag is a time tag or a location tag: Naïve scan and Scale-structure Identification. To simplify the discussion, we describe the methods only for time tags; the methods for location tags are similar.

3.1 Naïve Scan

Naïve scan is a standard algorithm in signal processing to detect a burst [4]. It divides the data's time distribution with scale value r and if the number of data in a segment is more than *Average(number of data in each segment)+ 2·Standard deviation(number of data in each segment)*, it determine that the segment is a burst.

Rattenbury et al. [3] gives the following method of using a Naïve scan to determine whether a tag is a time tag: For each segment i, let $T_r(x, i)$ be the usage count of tag x and $N_r(i) = \sum_x T_r(x, i)$. μN and σN respectively represent the average and standard deviation of $\left\{N_r(i)|i = 1 \cdots\right\}$. If $\frac{Tr(x,i)}{\mu N + 2\sigma N}$ is more than a certain threshold, the tag is determined to be a time tag.

3.2 Scale-structure Identification

The result of the Naïve scan method depends on the scale value r and segment separation determined by r. Hence, the selection of r determines whether a tag is a time tag or not.

To solve this scale problem, Rattenbury et al. presents a Scale-structure Identification method [3] that considers multiple scales based on Witkin's Scale-space method [5]. The Scale-structure Identification method defines scale values as follows: $R = \left\{r_k|k = 1 \cdot K, r_{k_1} > r_{k_2} \iff k_1 > k_2\right\}$, R is selected such that that $r_k = \alpha^k, \alpha > 1.1$.

For each scale value r, do the following:

1. In T_x , consider the graph over T_x where the edges of a node pair exist if the distance between the two nodes is less than r.
2. Consider the set of the connected graphs (the set of clusters) Y_r, and calculate the entropy $E_r = \sum_{Y \in Y_r} \frac{|Y|}{|T_x|} \log_2 \frac{|T_x|}{|Y|}$.

E_r indicates the degree to which the whole distribution is similar to one cluster (When $|Y_r| = 1$, $E_r = 0$).

If the time distribution T_x of x has more than one dense region for all scales, the probability that the distribution is similar to one cluster is high. In other words, for all scales, the value of Er is likely small when the tag is a time tag. Thus, the method determines that the tag is a Time Tag when E_r is less than a certain threshold.

Figure 1 shows how the clustering changes for different scale values.

4 Extracting the concepts of the time and location tags

The concepts of the time or location tag can be extracted by using the Naïve scan or a variant of Scale-structure Identification that we describe below. For the sake of brevity, we describe only the case of an identified time tag.

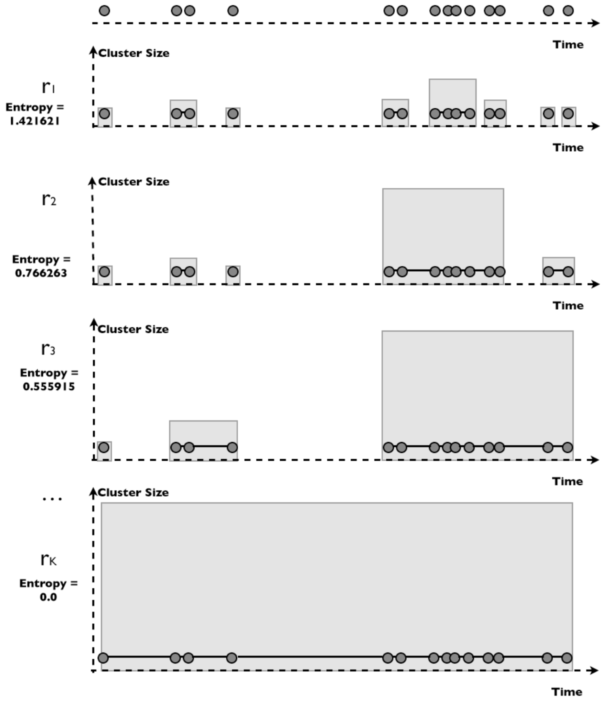


Fig. 1. An example of clustering

4.1 Naïve Scan

As described in Section 3, a Naïve scan determines that a segment is a burst if the number of data in the segment is more than a certain value. Hence, we can consider the segment is "the dense region" and the segment embodies the concept of a time tag.

4.2 Scale-structure Identification

The Naïve scan needs a pre-defined scale value and its results depend on this value: Thus, to get the best result, we divide a segment into multiple segments. To do so, we can employ a variant of Scale-structure Identification, which merges the results of analyzing clusters on all scale values. In order to extract the concepts of time tags, we have to solve two problems:

1. Which is the best clustering structure of scale value r to characterize the time distribution T_x ?
2. Once we know which clustering structure is the best, which cluster should be taken as embodying the concept of the time tag?

Selecting the clustering structure A better clustering structure has the following properties:

1. There is more than one cluster having a lot of nodes.
2. Clusters are distant from one another.

1 is needed because we want to specify the dense region of the distribution. If clusters are close, they could be combined. To represent the characteristics of the time distribution T_x, the Scale-structure Identification method calculates the entropy Er for each clustering structure, but entropy only indicates 1 not 2 because it does not consider the clusters' distance. Hence, we use compactness and isolation [6][7][8] as indications of 1 and 2. These measures are taken as evaluations of clusterings in the scale-space method:

$$compactness(Y_i) = \frac{\sum_{t \in Y_i} \exp^{-\|t-t_i\|^2/2r^2}}{\sum_{t \in Y_i} \sum_{Y_j \in Y} \exp^{-\|t-t_j\|^2/2r^2}}$$

$$isolation(Y_i) = \frac{\sum_{t \in Y_i} \exp^{-\|t-t_i\|^2/2r^2}}{\sum_{t} \exp^{-\|t-t_i\|^2/2r^2}}$$

t is each point in T_x , Y_i is the target cluster, Y is the set of tall clusters in the clustering structure, and t_i is the center point of Y_i . Compactness and isolation are related to a and b. Each value is less than 1.0, and bigger is better. Using the evaluation values of each cluster, we define evaluation formulas for the whole clustering structure [8]:

$$F_c(r) = \sum_{i}^{m} compactness(Y_i) - m$$

$$F_i(r) = \sum_{i}^{m} isolation(Y_i) - m$$

Here, m is the number of clusters in the clustering structure Y. The clustering structure is better if $F_c(r)$ and $F_i(r)$ are both bigger.

Characterizing the best clustering structure After selecting the best clustering structure, we determine which cluster best characterizes the clustering structure. If the number of nodes in a cluster is smaller than a percentage of all nodes, we consider that the cluster is characteristic; i.e., the cluster is the concept of the time tag. (Each tag can have multiple concepts.)

5 Evaluation

We implemented Naïve scan and Scale-structure Identification methods on actual Flickr data. We evaluated the following items:

time tags	august2007, carnival, comiket, cosmos, firework, fujirock, gameshow, ginkgo, gionmatsuri, jidai, june, newyear, obon, rama, september, snowboarding, tgs2006
location tags	beppu, chatan, daibutsu, enoshima, f1, hakodate, himeji, kanazawa, matsumoto, nara, otaru, roppongi, shinagawa, takayama, tokyodisneysea

Table 1. An example of time/location tags

1. Does the concept of the extracted time tag or location tag correspond to a concept identifiable by a human?
2. Does the concept extracted with Scale-structure Identification better correspond to human recognition than the concept extracted by the Naïve scan?
3. Is it possible to extract the tags' concept from the Flickr tag usage distributions?

5.1 Dataset

We collected photo data from Flickr, including photos taken between 2004/1/1 and 2007/12/31 and annotated with the location data of Japan. We excluded photo data whose upload date was before the taken date. On Flickr, the annotated location data to the photo was given an accuracy level between 1 and 16, with higher being more accurate. We removed photo data whose accuracy was less than 6.

We focused on tags annotating more than 100 photos and that were used by more than three users. The final data set contained 3,826,253 photos and 2,453 tags.

First, we manually determine that each tag is time tag or location tag and then randomly chose 100 tags from each set. Table 5.1 shows examples of time and location tags.

5.2 Experiment

We applied Naïve scan and Scale-structure Identification (SSI) to the selected tags. We analyzed multiple scales and selected the best scale by method and output the result by scale. We output the results of all scales to see if a suitable scale was chosen by the method. We use the scale values $r_k = 2, 4, \cdots$.

Each 100 time and location tags was manually ranked from 1 (low) to 5 (high). Figure 2 shows an example output, and Figure 3 shows an example rating.

The comparisons are:

- SSI: Scores for the results of the chosen scale by SSI.
- Best SSI: Scores for the best results of the SSI.
- Naïve Scan: Scores for the results of the Naïve scan.

The Naïve scan needed a pre-defined scale value so we calculated the average score for each scale and chose the best scale for it.

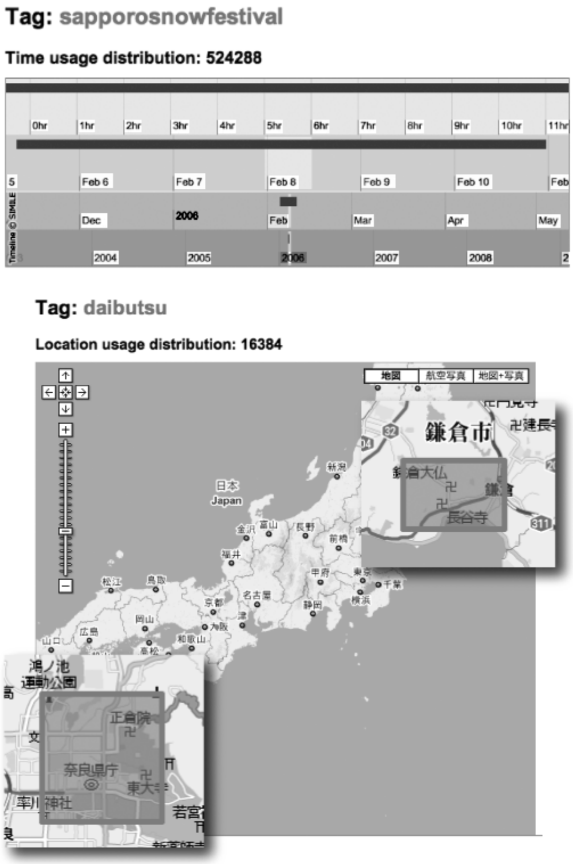


Fig. 2. An example of output

5.3 Result

Table 5.3 shows the average score for each method and the precision for right results when the score is higher than two. Figure 4 shows the distribution of differences between each method's scores and the best SSI scores. When the values are positive, the score for each method is better than the Best SSI. Lateral axis is the number of tags.

The average scores for SSI were near three, and precision was higher than 50%. These results confirm the concepts correspond to ones recognizable by humans. Average score and precision are higher for SSI than for Naïve scan. This confirms that our SSI method can get better results than the existing methods.

From Figure 4, it is often the case that the scores for Naïve scan are better (+1 to +4) or much worse (-2 to -4) than the best SSI scores; hence, the Naïve

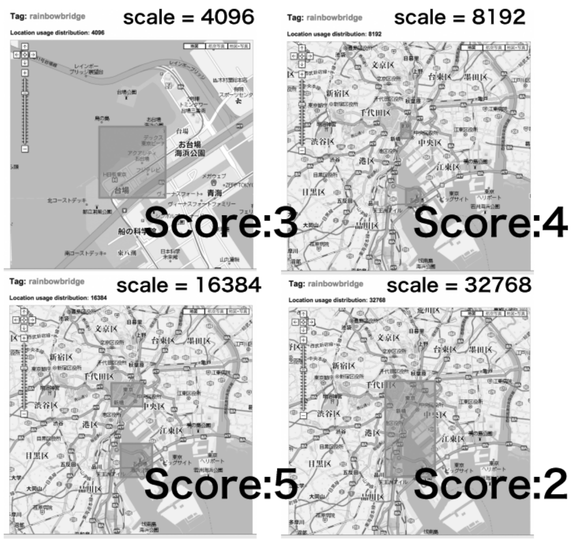


Fig. 3. An example of rating

		Average Score	Precision
Time	Best SSI	3.74	0.80
	SSI	2.70	0.56
	Naïve Scan	2.64	0.53
Location	Best SSI	3.86	0.84
	SSI	2.67	0.57
	Naïve Scan	2.62	0.49

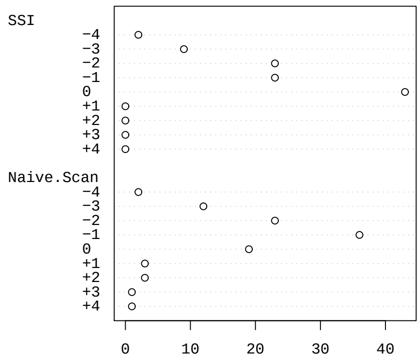
Table 2. Average score and precision for each method

scan's results vary in quality. Because SSI selects a suitable scale for each tag, its output is not as mismatched with human recognition as the Naïve method.

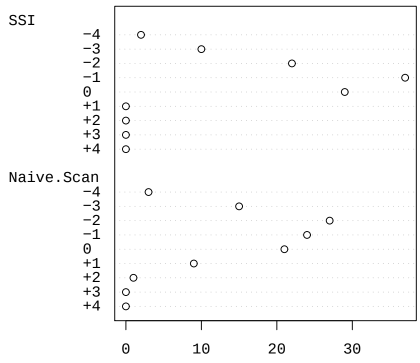
Therefore, the average best SSI score is near four and precision is higher than 80%. This means if we can select a suitable scale,we can extract the time or location concept of tags. Hence, our goal of extracting the concept of tags from their usage pattern is achievable.

6 Related Work

Researches about tagging system are well-practiced. Golder and Huberman [9] investigate how user do tagging and bookmarking on del.icio.us, a popular social bookmarking service. They discover regularities in user activity, tag frequencies and bursts of popularity in bookmarking. They also categorize the tags used for bookmarking into seven classes by their function. On the other hand, Marlow et



(a) Time tag



(b) Location tag

Fig. 4. Distribution of difference between each method scores and Best SSI scores.

al. [10] identify taxonomy of tagging systems' design and user incentives. Moreover, they showed dynamics of Flickr and del.icio.us systems are quite different. Another research [11] survey the motivations for annotation and tagging photographs in mobile and online media, especially Flickr and ZoneTag [12]. [13] analyze tagging behaviour of users on Flickr and show the distribution of Flickr tags over the most common WordNet categories.

Improving usefulness of tags has been of increasing research interest. There is a work to assign specific schemes to facilitate interoperability between tagging systems [14]. Some methods are proposed to derive semantic structures from tags: obtaining semantic relations between the tags by using online ontologies [15], cluster tags by organizing undirected graph from tag space (Each node corresponds to a tag and Each edge is weighted related to co-occurrence frequency of a tag pair) [16] and deriving hierarchical semantics of tags by using unsupervised model [17]. Xu et al. [18] introduce a tag suggestion system. This system spot high-quality tags which determined by their popularity, coverage, etc.

Some research analyze the tags on Flickr using temporally information [19] or spatially information [20]. The system described in [19] generate the interesting tags on Flickr during specific time period by computing interestingness of each tags and visualize them. World Explorer [20] also focuses representative tags for each spatial region and obtain them by using techniques of multi-level clustering and TF-IDF based scoring.

There have been some previous works to extract semantic from Flickr tags, directly related to our work. Scale-structure Identification [3], as applied here, determines whether tags on Flickr are time and/or location Tags. We also adapted it to extract the concepts of such tags. Schmitz et al. [21] tried to extract an ontology from Flickr tags using tag co-occurrence relations and organizing subsumption model and the research in [13] also analyzed the tag co-occurrence to build a tag recommendation system. Both works focuses on the tag semantics, similar to ours, but their approach is to extract synonyms from tag co-occurrence relations so the target tags' concept is different from ours.

7 Conclusion

In this paper, we focused on the new challenge of extracting the temporal and spatial concepts of tags from the tags' temporal/spatial usage distributions. To do so we modified the Scale-structure Identification, which is the existing method to determine whether a tag is related to a time or location, and showed that it can extract the temporal and spatial concepts of tags with higher accuracy than a Naïve scan in an experiment using actual Flickr data. We showed that if the method is able to select the suitable scale, it can extract the temporal or spatial concept of the tags with a high degree of accuracy, higher than 80%.

References

1. Vander Wal, T.: Folksonomy coinage and definition. http://www.vanderwal.net/folksonomy.html (2007)
2. : Flickr. http://www.flickr.com
3. Rattenbury, T., Good, N., Naaman, M.: Towards automatic extraction of event and place semantics from flickr tags. In: SIGIR '07: Proceedings of the 30th annual international ACM SIGIR conference on Research and development in information retrieval, New York, NY, USA, ACM (2007) 103–110
4. Vlachos, M., Meek, C., Vagena, Z., Gunopulos, D.: Identifying similarities, periodicities and bursts for online search queries. In: SIGMOD '04: Proceedings of the 2004 ACM SIGMOD international conference on Management of data, New York, NY, USA, ACM (2004) 131–142
5. Witkin, A.: Scale space filtering. Readings in Computer Vision: Issues, Problems, Principles, and Paradigms (1987)
6. Wong, Y.F.: Clustering data by melting. Neural Comput. **5**(1) (1993) 89–104
7. Chakravarthy, S.V., Ghosh, J.: Scale-based clustering using the radial basis function network. IEEE Transactions on Neural Networks **7**(5) (September 1996) 1250–1261
8. Leung, Y., Zhang, J.S., Xu, Z.B.: Clustering by scale-space filtering. IEEE Trans. Pattern Anal. Mach. Intell. **22**(12) (2000) 1396–1410
9. Golder, S., Huberman, B.: The Structure of Collaborative Tagging Systems. Arxiv preprint cs.DL/0508082 (2005)
10. Marlow, C., Naaman, M., Boyd, D., Davis, M.: HT06, tagging paper, taxonomy, Flickr, academic article, to read. Proceedings of the seventeenth conference on Hypertext and hypermedia (2006) 31–40
11. Ames, M., Naaman, M.: Why we tag: motivations for annotation in mobile and online media. Proceedings of the SIGCHI conference on Human factors in computing systems (2007) 971–980
12. : Zonetag. http://zonetag.research.yahoo.com/

13. Sigurbjörnsson, B., van Zwol, R.: Flickr tag recommendation based on collective knowledge. In: WWW '08: Proceeding of the 17th international conference on World Wide Web, New York, NY, USA, ACM (2008) 327–336
14. Gruber, T.: Ontology of folksonomy: A mash-up of apples and oranges (2005)
15. Angeletou, S., Sabou, M., Specia, L., Motta, E.: Bridging the Gap Between Folksonomies and the Semantic Web: An Experience Report. Bridging the Gap between Semantic Web and Web Workshop at ESWC2006, Budva, Montenegro **2** (2006)
16. Begelman, G., Keller, P., Smadja, F.: Automated Tag Clustering: Improving search and exploration in the tag space. Collaborative Web Tagging Workshop at WWW2006, Edinburgh, Scotland (2006)
17. Zhou, M., Bao, S., Wu, X., Yu, Y.: An Unsupervised Model for Exploring Hierarchical Semantics from Social Annotations. The Semantic Web, 6th International Semantic Web Conference, 2nd Asian Semantic Web Conference, ISWC (2007) 680–693
18. Xu, Z., Fu, Y., Mao, J., Su, D.: Towards the semantic web: Collaborative tag suggestions. Collaborative Web Tagging Workshop at WWW2006, Edinburgh, Scotland (2006)
19. Dubinko, M., Kumar, R., Magnani, J., Novak, J., Raghavan, P., Tomkins, A.: Visualizing tags over time. In: WWW '06: Proceedings of the 15th international conference on World Wide Web, New York, NY, USA, ACM (2006) 193–202
20. Ahern, S., Naaman, M., Nair, R., Yang, J.: World explorer: visualizing aggregate data from unstructured text in geo-referenced collections. Proceedings of the 2007 conference on Digital libraries (2007) 1–10
21. Schmitz, P.: Inducing ontology from flickr tags. Collaborative Web Tagging Workshop at WWW2006, Edinburgh, Scotland, May (2006)