Users Tagging Behavior and the Effect of Recommendation

Angelina Ziesemer
PUCRS
Faculdade de Informática
Porto Alegre – RS, Brazil
angelina.ziesemer@acad.pucrs.br

James Blustein
Dalhousie University
Faculty of Computer Science
Halifax – NS, Canada
jamie@cs.dal.ca

Milene Silveira
PUCRS
Faculdade de Informática
Porto Alegre – RS, Brazil
milene.silveira@pucrs.br

ABSTRACT
Users' tagging behavior has evolved. Formerly, the majority of users employed tags to label explicit content presented in resources. Now many users assign tags to express more than mere content description. The primary goal of this paper is to investigate how recommender systems change users' tagging behavior. We focused our investigation in the tagging task, modeling tags from a semiotic point of view. The results of a user experiment we conducted show that recommender systems can drive users to change the tag structure adopted.

ACM Classification Keywords
H.5.3 Information Interfaces and Presentation: Miscellaneous

Author Keywords
self-expression, tagging, recommendation, image-sharing, semiotic, user modeling

INTRODUCTION
As technology evolves, tags have become an important tool for content categorization and description. In general, the taxonomic categorization of subjects or things is made by experts, but on the Web, social tagging is a popular approach used by regular (non-expert) users and it enables users to index content, improving the retrieval results from images, videos, etc. However, tagging is a repetitive and tedious task and users may not know how and why they should use tags, leading designers to resort to approaches and tools that can improve and encourage users to assign tags.

Social media network users have expanded the type of tags assigned to express more than the explicit content in the resources[1]. Users have distinct motivations for tagging, such as self representation and opinion expression, which change the type of tag adopted [3] from tags with commonly accepted meanings to personal tags. This behavior expand the common vocabulary used for categorization, which influences also how such tags are composed [5]. By using tags on images users can express not only a description of the ostensive and rarely disputable aspects of the image, but also the concepts the image evokes to the user who tags.

In order to facilitate and improve the tagging task and support the use of quality tags, recommender systems have arisen to help users choose putative tags. In this work, we are interested in the effect of recommender systems on the tag structure assigned by users, instead of the class of tag users assign. We focuses tagging behavior on the tag structure (we define tags as linguistic signs in the Semiotic sense [2], words as units or sequences to represent a tag), adopted before and after the recommendation aid. We conducted an experiment asking participants to assign tags to photos with and without the assistance of recommendation. We also used a demographic survey to gather information about participant profiles. We posed the following research questions to help us understand the role recommender systems play in tagging behavior:

- RQ1: Does the use of a recommender system change the tagging behavior regarding structure of tags users employ?
- RQ2: Does the tagging behavior change according to the class of photo being tagged, regardless of the system used?

Our findings show that recommendation can drive users to change the structure of tags. Results can benefit designers by providing guidance as to approaches and data sets which are appropriate to their system goals and user's needs. Instead of considering tags as noise based on their structure and popularity, designers could use the tagging semiotic perspective we proposed as tool for determining users' profiles based on their tagging behavior and manage recommendation for social communication or data description.

METHOD
Since we had two main conditions to investigate (tagging with and without the support of recommendation), a total of 57 participants (26 female and 31 male, with a mean age of 26 years old) were divided in two groups that had the experimental conditions changed: group one (G1) was asked to assigned tags to photos without the recommendation support and after they assigned tags to the other set of photos in a different order supported by a recommender system. Group G2 was exposed to the same conditions in the opposite order.

[1]By "resources" we mean any online content that can be tagged, for example images, websites, documents, sound files.

IHC '16, October 04-07, 2016, São Paulo, Brazil
 ISBN 978-1-4503-5235-2/16/10…$15.00
DOI: http://dx.doi.org/10.1145/3033701.3033738

Also, participants were required to answer a demographic survey.

This study had a browser-based platform and a total of 7 photos publicly available on Google Images were presented to each participant: four images were present in both NR (no recommendation stage) and RS (recommender system stage) stages for the purpose of comparing behaviour. The other three images were presented only in the RS stage to address whether no previous experience with the image have difference on tagging language.

Content-oriented classification of photos

To avoid biases we randomly recruited fourteen individuals (that were not part of the main study reported in this paper) to conduct the photo content classification for the photos used in this study. Each participant received the photos in a random sequence and were asked to classify the content of each image by its level of information presented regarding context: situation (whether the concept represented in the photo stands out) and location awareness (if the location where the photo was taken is obvious in the photo content). They also classified whether important parts of the image (the content) was clearly delineated from both the foreground and background. Table 1 shows the photos classification reported as having high level of content regarding the classes previously described.

Table 1: Results from content-oriented classification of photos.

	Photos						
Classification	P_a	P_b	P_c	P_d	P_e	P_f	P_g
Prominence							
Background						×	
Foreground	×	×					×
Both			×	×	×		
Context (can be determined)							
Location					×	×	
Situation		×	×				

Tagging using Recommendation

As a design platform, we used a model [1] that has as its primary purpose the recommendation of tags. We chose this model for its simplicity and also, because it uses reference tags from users to recommend other tags (the so-called semi-automatic approach). In other words, after a participant assigns a tag he/she receives as recommendation a list of other tags that could be assigned to the same image. We instructed users how it works and the options they could select after typing a reference tag:

"Each time you type a tag, the system will recommend other tags. You can select the tags that you think are appropriate to the content being tagged".

We used a training dataset from Flickr for recommending tags with more than $600,000$ tags. The utility of tags is computed by the combination of three measures for later combination to present a ranking of tags that can fit with a reference tag[2].

[2]A tag typed by the users to the photo being tagged.

The recommendation model define each posting P_i as a triple $P_i = \langle u_i, r_i, T_i \rangle$ where $T_i = \{t_1, t_2 \ldots t_n\}$ is a set of tags assigned to resource r_i posted by user u_i. This approach uses a reference tag t to get similar tags based on its co-occurrence in $P(t) = \{P_i | t \in T_i\}$.

Initially, it computes the k-tags with the largest co-occurrence from $P(t)$. A function records the existence of t in T and it is used to rank the co-occurring tags t_j by $ranking(t, t_j) = \sum_{P_i \in P(t)} (t_j \subset T_i)$.

After, three measures are computed to punish (these measure takes from the top of the ranking tags used by few users but that are very frequent in the data set) those tags that are not relevant or popular. They are the co-occurrence, relevance and popularity measures as following described.

Co-occurrence $coo(t, t_j)$*:* this measure is a normalization of the previous ranking. It computes both t and t_j by the number of items that have t, resulting in a value that can range from 0 to 1 for each t_j.

Relevance $rel(t, t_j)$: this measure takes from the top of the ranking those tags that do not represent the community vocabulary, i.e. name of the resource owner, tag reference to personal content etc. It computes the number of users that used t and t_j by the number of items that have being found in the previous ranking by $ranking(t, t_j)$.

Popularity $pop(t, t_j)$*:* this measure computes the popularity of t_j, that is, how important t_j is to the set of users that use t. The popularity is related to the frequency of use of t_j by the community. This measure uses the conditional probability as bases for computing the number of users using tags t and t_j divided by the number of users that use t.

Finally a ranking of recommended tags is computed by the geometric mean using the three previous measures.

Processing of Tagging Dimensions

In this work, we are interested in the use of words as units or sequences to represent a tag. To model these differences we approach tagging from a semiotic point of view. A language is a system of signs that can express ideas to represent something for someone (the user in our case) [4]. To prepare a set of tags to assign to a resource, users have to choose a tag (sign) from amongst many possible tags; The tags that are assigned give meaning to the content being tagged.

The meaning of signs arises from the differences between signifiers, and these differences are syntagmatic (concerning the relative position of words in a text, a combination of signifiers that aims to form a syntactically-correct meaningful text) and paradigmatic (concerning substitution, associated signifiers each one differing significantly) [2, 4]. In language, a sentence is a syntagm of words, a syntactic construction, and by analogy tags as sentence have a syntagmatic structure. Any units or elements of language presented in sequence can be represented by a chain [2]. Figure 1 shows the axes of the structure and relation of paradigms and syntagms.

The paradigmatic axis represents units assigned as tags. In general, these are putative tags used to describe objects, places,

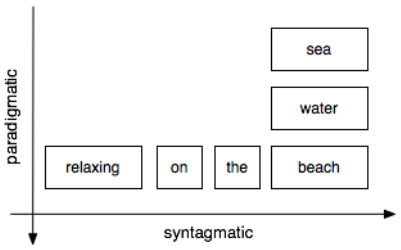

Figure 1: The dimensions of syntagmatic and paradigmatic relations.

people. These are tags that most people would agree to assign. These type of tags are helpful to describe the content of the photo [3]. On the other hand syntagmatic tags have a distinct structure, such as, *"Just saying", "Living my life"* and users in general assign it to express more than a description of the resource's explicit content. According to the definition of structure based on the Saussure chain, we use distinctive processes to quantify the tag structures in each stage of this study.

The tags from the NR stage were manually coded as paradigmatic or syntagmatic once there was no difference among their source (all of them were added without the recommender system assistance). In the RS stage, we coded tags using activity logs collected comprising information of all tags assigned, each one with an association about the original source (tags added by users as reference tags to get recommendation or tags recommended to users assign). This step in the process allows for the observation and comparison of the frequency of reference tags against those tags that were recommended. We observed the long tail of the power-law distribution of tags gathered in this study to classify the structure of each tag and, as we expected, the majority of syntagmatic tags were in the long tail. After the classification and frequency computation of each tag structure from both stages of the experiment, we focused our analysis in the research questions previously stated using statistical methods for each distinct situation.

RESULTS

Type of System vs. Tag Structure

In this work we do not argue that paradigmatic tags are better than syntagmatic tags. Instead, we aim to verify if the users tagging behavior, regarding the tag structure, changes once they receive tag recommendation. To address our first research question, we hypothesize that there is no relationship among the type of system used and the tag structure adopted. We compared the tags assigned to the same set of photos ($P_{a,b,c,d}$) presented in both stages of the experiment which resulted in a total of 823 tags in the RS stage for both G1 and G2. Table 2 summarizes the tag structures found. Results show that the proportion of syntagmatic tags changed when participants were aided by the recommender system. Results of a χ^2 test indicate an association among the variables for both groups. Next, we investigated whether this behavior also changes according to the class of photo being tagged.

Tag Structure	G1		G2	
	NR	RS	NR	RS
Paradigmatic	479 (72%)	439 (94%)	355 (73%)	324 (91%)
Syntagmatic	193 (28%)	29 (6%)	132 (27%)	32 (9%)
Total	672	467	487	356

Table 2: Classification of tags structure for the photos presented in both stage for each group. The p-value ($p < 0.01$) shows that there is an association among the type of system used and the type of tag structure assigned.

Photo Class vs. Tag Structure

The photos chosen for this experiment highlight their content position, the context they represent (situation, concept or message standing out in the photo) and the context regarding the location where the photos were taken. For this analysis, we expect that the proportion of syntagmatic and paradigmatic tags does not change regardless of the photo class.

Table 3 shows the proportion of tags assigned in both stages (NR and RS). We found much evidence that photos with high associated context/situation are related to syntagmatic tags. In the NR stage, we found, mainly for the G1, that the proportion of syntagmatic tags does not occur with the same proportion of paradigmatic tags to all photos but P_b. Photo P_b stands out with its context and foreground objects. According to the p-value resulted from the z-test of proportion, its (P_b) proportion of syntagmatic and paradigmatic tags does not change significantly ($p > 0.05$). Differently, all the other photos presented a significantly result regarding the proportion of syntagmatic tags ($p < 0.01$) even those photos ($P_{e,f,g}$) that were presented only in the RS stage still showed the same tagging behavior of photos that were in both stages, which shows that the previous visualization of photos did not influence the tagging task in this study.

	Stage	Tag	P_a	P_b	P_c	P_d	P_e	P_f	P_g
G1	NR	Syntag.	0.29	0.45	0.20	0.23	-	-	-
		Parad.	0.71	0.55	0.80	0.77	-	-	-
	RS	Syntag.	0.04	0.12	0.02	0.05	0.04	0.05	0.04
		Parad.	0.96	0.88	0.98	0.95	0.96	0.95	0.96
G2	NR	Syntag.	0.32	0.39	0.16	0.23	-	-	-
		Parad.	0.68	0.61	0.84	0.77	-	-	-
	RS	Syntag.	0.07	0.18	0.07	0.05	0.02	0.03	0.06
		Parad.	0.93	0.82	0.93	0.95	0.98	0.97	0.94

Table 3: Proportion of syntagmatic and paradigmatic tags in both stages of the experiment to each photo.

We performed the same test to compare the tag structure proportion from one stage to another. When participants were aided by recommendation their tagging behavior changed ($p < 0.01$) to the photos presented in both stages. This effect occurred also to (P_b), which presented, mainly for G1, no difference in proportion of syntagmatic and paradigmatic tags.

Tag Structure vs. Type of Tag

To illustrate the differences among the vocabulary agreement on both stages and the type of tag assigned, Figure 2 (A) shows the distribution of tags in the NR vs. RS stage, and

(B) the distribution of reference and recommended tags in the RS stage. We extracted the tags that were in the head and in the long-tail of the power-law distribution of tags. First, we looked to the tags more frequently assigned in the NR stage (A). The head of the power-law, that represent the common vocabulary of participants, represented 58% of tags assigned in this stage. Most frequently assigned tags had paradigmatic structure and were used to describe the photo content as *beach, cat, cute, metro, friends*. In the head of the power-law, 6% of tags had syntagmatic structure. In the other hand, in the long tail, represented by 42% of tags that were assigned only once, 57% of them had syntagmatic structure. The syntagmatic tags found seem to be motivated by social communication, self-expression (opinions, emotions) and personal tags: *happy-monday, funnyday, crazyexperience*, *nosensefriend*[sic], *mypet*. Comparing the power-law distribution from the RS stage, tags assigned had a higher agreement than the NR stage. The head of the power-law of the RS stage represents 77% of tags and the long tail 23%.

Looking to the power-law distribution of the type of tags of the RS stage (Figure 2 (B)), we found that 62% of tags were assigned as reference tags and 48% were assigned based on recommendation. From the set of reference tags 90% were paradigmatic and 10% syntagmatic on which 87% of them were in the long tail: *supercute, dontfallasleepatthesubway*. When we look to list of recommended tags assigned, syntagmatic tags represent only 2% of tags recommended and 65% of them were in the long tail: *morningafter, familyvacation*. Syntagmatic tags in the head of the power-law in the RS stage seem to be more related to content description (*sunnyday, blueeyes*) while in the NR stage, besides the syntagmatic tags in the head used to content description *whitecat, bluesea*, were also found tags for social communication and related to the photo context (*loveit, bestfriends, bestpicture*). This tagging behavior suggest that may there is difference on syntagmatic tags that are from commonly agreement from those that are not. This observation opens space for future investigation since the quantity of syntagmatic tags resulted from our study is not enough to generalize a conclusion for the vocabulary commonly used for this type of tag.

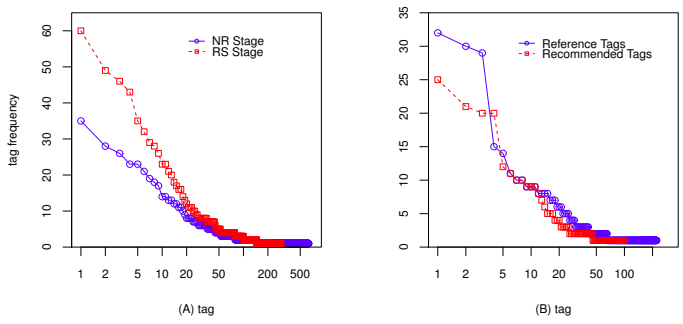

Figure 2: Power-law distribution comparing NR vs. RS stage (A) and the distribution of reference and recommended tags (B).

DISCUSSION AND CONCLUSION

Understanding how users perform the same task in different environments can provide insight for designers to decide among distinct approaches according to users and system needs. This work investigated the potential of recommender system on change tagging behavior from a semiotic point of view.

Based on the results we found for RQ1, participants more frequently assigned syntagmatic tags in the NR stage and this behavior changed in the RS stage. The recommender approach used in this experiment worked in a interesting way that improved the number of paradigmatic tags assigned which also improve the homogeneity not only of tags distribution but also of tag structure. An important point is that although participants changed their tagging behavior in the RS stage, syntagmatic tags were also assigned supported by recommendation. This behavior calls for investigation into what distinguishes syntagmatic tags that fall in pool of commonly regarded tags, and are more likely to be recommended, since most of them were presented in the long tail for both stages.

Moreover, we found an association between the tag structure used and the type of photo being tagged (RQ2). This behavior was observed in the NR stage, syntagmatic tags were mainly used by participants to express the context of the photo regarding its situation, a type of social communication while assigning tags. This behavior changed when participants assigned tags in the RS stage: more paradigmatic tags were assigned compared to the proportion of tags found in the NR stage. So, regardless of the class of photos being tagged participants adopted more paradigmatic tags to all photos during the RS stage.

As potential to improve tagging, this work can encourage designers to model the semiotic tagging behavior before choosing any recommendation approach. It can help designer to identify users motivation (tagging for social communication, indexing, categorization) for tagging, before adopting any approach for recommendation.

REFERENCES

1. Angelina de CA Ziesemer and Joao Batista S de Oliveira. 2013. Keep Querying and Tag on: Collaborative Folksonomy Using Model-Based Recommendation. In *International Conference on Collaboration and Technology*. Springer, 10–17.
2. Ferdinand De Saussure. 1916; 2011. *Course in general linguistics*. Columbia University Press.
3. Manish Gupta, Rui Li, Zhijun Yin, and Jiawei Han. 2010. Survey on Social Tagging Techniques. *SIGKDD Explor. Newsl.* 12, 1 (Nov. 2010), 58–72.
4. Kaja Silverman. 1984. *The subject of semiotics*. Oxford University Press.
5. Csaba Veres. 2006. The language of folksonomies: What tags reveal about user classification. In *Natural language processing and information systems*. Springer, 58–69.