Multilingual Tagging Behaviour:
The role of recommender systems

Angelina Ziesemer
PUCRS
Faculdade de Informática
Av. Ipiranga 6681, Prédio 32
Porto Alegre - RS, Brazil

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

Milene Selbach Silveira
PUCRS
Faculdade de Informática
Av. Ipiranga 6681, Prédio 32
Porto Alegre - RS, Brazil
milene.silveira@pucrs.br

ABSTRACT

This paper investigates how recommendation affects tagging behaviour regarding the language adopted in tags. We conducted a study to compare the tags assigned to digital images with and without the support of a recommender system. Results pointed to have an association between users' language used to assign tags and the type of systems supporting this task.

Keywords

tagging, recommendation, multilingual, digital image, annotation

1. INTRODUCTION

Tag recommender systems have arisen to help users choose tags most suited to lead to better (more accurate, more efficient, more satisfying, etc.) content retrieval. One popular approach used for such recommendations is to suggest tags which co-occur [6, 5, 1]. When users assign recommended tags the set of their tags becomes more homogeneous [1]. However, what has not yet been studied is whether the resulting homogeneity of tags also applies to the language adopted for tagging.

Understanding how users perform the same task in different environments can provide insight for designers to decide among distinct approaches according to user and system needs. The language used for tagging has several implications for the dissemination of content on the Web [4]. In general, co-occurrence based approaches for recommending tags do not take in account the user's language, but rather use the collection of tags that co-occur with a target tag to recommend other tags.

In this work we report a study investigating whether the presence of recommendation on tagging system can change the users' choices of the natural language they use for tagging digital images. Participants from whom tags were collected were residents of a Portuguese-speaking country and no particular previous definition regarding the language for tagging was given. The results show that there is difference in the language adopted to assign tags when users are supported by recommendation.

2. METHOD

Because we had two experimental conditions to test (tagging with and without the recommendation aid), a counterbalancing design approach was employed in the current study. A total of 57 participants, all of them residents in a Portuguese-speaker country, were partitioned in two groups: G1 had 33 participants (16 female, 17 male, with a mean age of 27 years); G2 had 24 participants (10 female and 14 male, with a mean age of 25). Participants were presented the same images, interfaces, and were asked to perform the same tasks, but each group received the conditions in reverse order. As a design platform for recommendation, we used a model [1] primarily intended to recommend tags based on co-occurrence. This approach computes the utility of tags using a combination of three measures to compute a ranking of similar tags based on a *reference tag* (a tag the user assigned to the image before receiving any recommendations). As source for recommending tags we used a training dataset from Flickr composed by more than $600,000$ tags. Participants had to assign tags to seven distinct images publicly available on Google Images. As is common in such research, images were classified by content [2]. Four images were present in both NR and RS 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.

2.1 Classifying Languages

To process the language of tags assigned in this experiment, we use a standalone language identification tool based on a Naïve Bayes classifier [3]. This classifier provides a probability estimate of the natural language from which a given set of words are drawn. By performing the language identification and observing a bunch of tags and its probability estimation resulted, we found that some users tagged images multilingually, so the language classifier was useful to estimate a language score for each image classifying it as mainly assigned with tags in English (EN) or Portuguese (PT), the two main languages used by participants in the tagging task.

3. FINDINGS

Our first analysis assessed the effect of using, or not using, a recommender system on the language used to assign tags. The set of tags assigned to each image by each participant was classified as either PT or EN. Table 1 shows the difference in the proportion of images and the main language used by each group to images that were presented in both stages of this study.

Table 1: Comparing the proportion of images assigned mainly in PT and EN in both stage of the experiments.

	G1		G2	
	NR	RS	NR	RS
PT	81 (61%)	24 (18%)	63 (65%)	25 (26%)
EN	51 (39%)	108 (82%)	33 (35%)	71 (74%)

When not using the recommender system, participants tagged fewer images using EN (G1: mean = 1.54 SD = 1.60; G2: mean = 1.37, SD = 1.71). However in the RS stage, more images were tagged mainly in EN (G1: mean = 3.27, SD = 1.30; G2: mean = 2.91, SD = 1.28). A (paired) Wilcoxon signed-rank test indicated that the mean of images with tags in EN changed ($p < 0.01$) from one stage to another for both groups. This behaviour also was found when we looked to the language used in each image individually, before and after recommendation (McNemar $p < 0.01$) and also for the images that were tagged only in the RS stage.

To make sure that the results found in this study were not narrowed by a few participants' behaviour, we looked to their results individually. We classified users as: PT-taggers, EN-taggers or multilingual-taggers (ML); PT-taggers — had all their images classified mainly by tags assigned in PT; and ML-taggers had a mix of images tagged in EN and PT. Figure 1 shows the proportion of participants and their respective tagging behaviour in each stage of this study.

At the NR stage, 45% of participants were classified as PT taggers. However, this behaviour changed in the RS stage, only 8% of them kept tagging images mainly with tags in PT. In the RS stage, the majority of PT taggers switched their tagging language and behaved as EN- and ML-taggers.
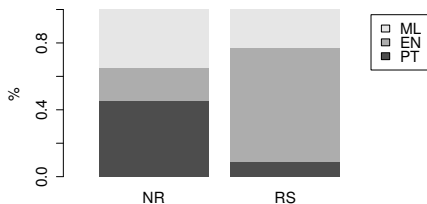

Figure 1: Comparing participants' language chosen for tagging in both NR and RS stages.

To try understand participants' behaviour, we examined the order of tags assigned in the RS stage: We noticed that at first some images received reference tags in PT but the following reference tags were assigned in EN. We hypothesized that, as participants received tag recommendation in EN, they switched the language of reference tags. However, individual users' behaviour needs future investigation.

4. DISCUSSION AND CONCLUSION

The combination of semi-automatic recommendation and tags' co-occurrence is an interesting and important approach used for recommendation, in part because the users' reference tags are the seed for recommending other tags. This approach is useful to make rich annotation, and also intends to improve the user experience by decreasing the effort to assign tags while still allowing users to use personal tags. In our study we have shown that the recommendation approach used affects the language choice for tagging. Results indicate that the quantity of images assigned mainly with tags in EN changed in the RS stage (compared to the NR stage). Because co-occurrence based approaches use the link among tags to recommend other tags, the results found here have several implications for the design process of tagging recommender systems. The approach increased the language homogeneity (EN) of tags which could result in a cultural isolation of online indexed content. We are aware that the training dataset used for recommendation has a representative quantity of tags in EN and consequently many tags in EN can co-occur with tags in another languages. However, the social tag dataset we used represents the natural imbalance of the language on the Web and how tags are connected, which reinforces that the EN language functions as a hub to other languages. On the other hand, the interface used in the RS stage only recommends tags based on the reference tags assigned, so users still had the autonomy to assign their own tags and keep their vocabulary without any recommended tag. The lack of switching (or converting) languages is evident when one considers that even in the RS stage some users continued as PT-taggers. These findings highlight the needed to investigate whether there is distinct users profile that are more likely to accept multilingual recommendations.

5. REFERENCES

[1] A. de CA Ziesemer and J. B. S. de Oliveira. Keep querying and tag on: Collaborative folksonomy using model-based recommendation. In *Collaboration and Technology*, pages 10–17. Springer, 2013.

[2] W. Dong and W.-T. Fu. Cultural difference in image tagging. In *Proc. SIGCHI*, pages 981–984. ACM, 2010.

[3] M. Lui and T. Baldwin. Cross-domain feature selection for language identification. In *Intl. Joint Conf. on Natural Language Processing*, pages 553–561, 2011.

[4] S. Ronen, B. Gonçalves, K. Z. Hu, A. Vespignani, S. Pinker, and C. A. Hidalgo. Links that speak: The global language network and its association with global fame. *Proc. of the Nat'l Acad. of Sciences*, 111(52):E5616–E5622, 2014.

[5] B. Sigurbjörnsson and R. van Zwol. Flickr tag recommendation based on collective knowledge. In *Proc. WWW*, pages 327–336, 2008.

[6] C. Wartena, R. Brussee, and M. Wibbels. Using tag co-occurrence for recommendation. In *Intelligent Systems Design and Applications, 2009. ISDA'09. Ninth International Conference on*, pages 273–278. IEEE, 2009.