Image Tagging and Search – A Gender Oriented Study

Adrian Popescu
Institut TELECOM/TELECOM Bretagne
Brest, France
adrian.popescu@telecom-bretagne.eu

Gregory Grefenstette
Exalead
Paris, France
gregory.grefenstette@exalead.com

ABSTRACT

Social computing sites constitute a valuable source of user-generated content for user modeling. Whereas user generated content and the mining of such content are well studied, little attention has been given in the literature to modeling the relationship between users' personal information and content. Here we analyze the relation of user gender to the choice of tags to describe a photo. A large user sample is examined to produce gender-related tagging vocabularies and tag representations. 1000 salient tags' male and female representations are compared and results indicate that there are important differences between gender based term choices in a large majority of cases. To test the influence of gender on retrieval, we built a gender sensitive image search prototype and tested it, using a survey. Results show that around two thirds of participants tend to prefer image search results obtained using tag representations of their own gender and that a third of participants have a clear preference for their own gender's results.

Categories and Subject Descriptors

H.3.3 [**Information Storage and Retrieval**]: Information Search and Retrieval—*search process*

General Terms

Algorithms, Experimentation, Human Factors

Keywords

Gender, social networks, Flickr, tags, image, search

1. INTRODUCTION

User selected tags are widely used to describe online content in images (Flickr, Picasa), in video (YouTube, Dailymotion) or in blogs (Technocrati). Beyond their utilitarian purpose of making retrieval possible, tag choice reveals implicit user point of view. Large scale tagging data has been studied extensively, but data about users' personal information has often been ignored in these published studies. Beyond mercenary considerations such as targeted advertising, awareness of personal information (such as gender) might help service providers improve user satisfaction in interaction with online services. Exploring this topic, we examine here the role of gender in tag production and image retrieval. We gathered textual metadata for a sample of 12,000 Flickr users (50% of each gender) who have declared their gender, from six different user-declared countries. We extracted gender-related tagging vocabularies and analyzed the distributions of tags in female and male vocabularies, over a sample of 1000 frequent tags. We exploited these vocabularies to develop a gender sensitive image search engine, which performed automatic query expansion using female and male related tags to original Flickr results. We assessed user satisfaction by surveys.

2. RELATED WORK

Efforts to understand the tagging process have generated a large body of research. Ames and Naaman [2] examined motivations for tagging and reported that people tag pictures both for personal and social purposes. Nov et al. [8] examined different factors that influence photo sharing. They showed that photo sharing becomes more selective over time and that there is a negative correlation between sharing habits and time. These cited studies focus on why and how people tag. What is tagged is also broadly studied and results are exploited, among other applications, for tag recommendation [14], [17] or folksonomy construction [13], [9]. In [14] and [17], the authors analyze a large body of Flickr tags to find tag correlations. Correlations are then used to suggest new tags based on already existing photo tags and their correlations to other tags. The authors of [14] also map tags into WordNet to extract the distribution of Flickr tags in different conceptual domains and report that main tag categories include artifacts and places. In addition to tag correlation, Wu et al. [17] also analyze image content in order to improve tag recommendation. The authors of [10] introduced a spike detection method for detecting events and places in a large corpus of geotagged Flickr metadata. ZoneTag [1] is a mobile application for uploading tagged images from a camera phone to Flickr which exploits the method introduced in [14]. In addition to the tags introduced by the user, ZoneTag proposes location and event related tags based on the users' tagging history, current coordinates and time stamp. The tagging process in ZoneTag is semiautomatic because the system only proposes tags and it is up to

WSM 2010 October 25, 2010, Firenze, Italy

the user to choose relevant tags. This is also true for the other presented tag recommendation methods and, to date, there exists no truly usable automatic image annotation system.

Schmitz [13] applies a statistical subsumption model to induce hierarchies of tags. The authors of [9] use statistical methods to build community-driven folksonomies and also show how to evaluate folksonomies by automatic comparison to a manually created taxonomy, demonstrating that meaningful relations can be extracted from large tag sets. Personalization is another exploitation of tag mining. Garg and Weber [6] present an interactive tag recommendation system driven by a user's tagging history. Here recommendation is user-centric whereas in [14] and [17] it was considered as a generic process.

Tag clustering has received much attention and is currently implemented in the Flickr interface. Following a query, search can be refined by selecting sets of related tags. In [15], the authors use Minimum Length Distance to find interesting groups of related pictures associated to a query. Their approach basically turns a clustering problem into a data compression problem and exploits both tags and URLs instead of tags only. The main roles of tag clustering or compression are to maximize results diversity and to surface representative images. Visual clustering [16] represents another way to diversify search results. In [16], clustering is applied to both ambiguous and non-ambiguous queries and is evaluated against manually clustered search results. Tests show that the approach tends to reproduce manual clustering in a majority of cases. However, increasing diversity is often synonymous to precision loss and a compromise needs to be found between these two characteristics of results [4].

The relation between authors' gender and texts is well studied for running text. [3] provides a detailed analysis of gender identification research over written documents, and propose a method for automatic gender identification based on text characteristics. Models of "female" and "male" texts are learned from lexical distributions and new texts can be classified based on these models. Nouns were found not to be discriminative for gender identification whereas adjectives were more often used by females. Also according to [3] some stop words are used more by one sex than the other. The authors of [7] analyzed genre and gender in weblogs and concluded, on the other hand, that language differences are more dependent on the type of blog entry (personal or news report) than gender. Rich [11] showed that women have a larger color vocabulary than men. This finding is coherent with the fact that texts authored by women tend to be more personal than those authored by men [3]. To the best of our knowledge, there exists no extensive study of the role of gender in tagging systems.

3. DATASET & PRELIMINARY ANALYSIS

One of our research objectives is the comparison of female and male users' tag choices. Using profile pages, we build a balanced user sample, containing 6000 females and 6000 males, with 1000 users of each gender from each of the following countries: USA, UK, France, Germany, Italy, and Spain. Contributors from different countries are selected in order to simulate the diversity of origins of Flickr users. We start with an initial list of 10,000 users and select those who disclose their location (situated in one of the six countries) and gender. This list is enriched with new users picked from the list of contacts of users already in the list until 1000 users of each gender are selected for each country. Next, for each user, we download textual metadata using the Flickr API and analyze these metadata in order to extract a list of unique tags used. After this step we can report some preliminary findings about the user sample. Americans upload the largest number of images on Flickr (over 1800), followed by British users (around 1450). Contributors from other countries (FR, DE, IT, ES) have significantly smaller contributions, with average portfolio volumes under 50% of American portfolios. With the exception of Americans, males upload more images than females, a tendency which is most obvious for Germans (995 vs. 607 photos on average) and Spanish (767 vs. 441 photos). Except for Italians, who tag in a detailed manner, the average number of unique tags used is well correlated with average contributions. It is largest for Americans (over 1000), followed by British (over 900) and is smallest for Germans and Spanish (just over 400 tags). Average individual photo tagging detail is higher for men in all cases, with the largest difference for Americans (5.2 tags/photo for women and 6.8 for men). The average size of tagging vocabularies is larger for men than for women, with the exception of Italians (591–F vs. 569–M). The difference is particularly evident for Spanish users (472-M vs. 350-F unique tags).

Table 1: Top 10 tags in the gender vocabularies by number of different users having annotated images with tags.

Gender	Tags
Female	red, blue, green, flower, sky, flowers, water, tree, white, pink
Male	sunset, sky, night, water, red, beach, blue, tree, portrait, light

4. TAGGING VOCABULARIES

Language variation with gender has been studied for formal texts [3] and blogs [7] but not for tagging systems, such as Flickr. We analyzed the tags provided by the 12,000 selected users in order to create gender vocabularies. The importance of a tag in a gender vocabulary is given by the number of different users who annotated photos with the respective term. This simple tag popularity measure is inspired by the landmark popularity score described in [5]. The measure is suited for expressing the importance of a tag in a community because it eliminates bias introduced by bulk tagging (individual users who tag large volumes of photos with the same terms).

We compute the top 1000 female and male tags by number of people who used the tag, presenting the most frequent 5 tags for each gender in table 1. As expected, all top tags are general words and can be associated to a large number of pictures. 6 out of the top 20 tags are unique to one gender and this is a first indication that there are notable differences between top things tagged by women and men. Moreover, tag ranking in table 1 is quite different for the two genders. For instance, color tags are much more important for women than for men. Top three female tags are colors, whereas the first color is ranked only 5th in the male vocabulary. This result shows that not only women use of larger color vocabularies [11] but they use these terms more frequently.

Table 2: Tags by gender predominance score. Top 5 tags in each score range are presented.

Score	Female predom.	Male predom.
> 2	doll, strawberry, kitty, bunny, tulips	panorama, hdr, longexposure, leaf, skyscraper
1.75−2	heart, chocolate, cats, roses, sweet	flash, police, sport, railway, auto
1.5 − 1.75	pink, purple, shoes, feet, fdsflickrtoys	tower, usa, nikon, truck, war
1.25 − 1.5	spring, love, rose, leaves, baby	bridge, car, train, moon, architecture
1−1.25	red, blue, green, flower, tree	sky, sunset, water, beach, night

To examine differences between the gender vocabularies beyond top elements, we compute a predominance score – the ratio between the two gender related usages of a tag. This operation is performed on the top 1000 tags obtained over the entire user sample. We present the most salient tags for different predominance score ranges in table 2. Scores for female predominance correspond to the female/male users ratio and scores for male predominance to male/female users. Only one tag (bird) has the same female and male score, 423 are predominantly female and 576 predominantly male. Overall, 56% of tags are in the first score range (1 – 1.25) and have weak gender orientation. 15% have predominance score higher than 1.5, meaning that they are much more used by one gender. Results in table 2 indicate that predominantly female tags are often personal whereas predominantly male tags are more neutral or technical.

5. MUTUAL TAG REPRESENTATION

For each tag, we extract a list of related tags and then compute a similarity score between female and male representations. One solution to find related terms is use a co-occurrence model (i.e. count the number of different users that tag with pairs of terms). However, initial tests showed that the method tends to discover too many general terms, which are associated to a large number of different tags. This type of extraction is well illustrated by "Tag Cluster" presented on Flickr pages after a query. We introduce a mutual information-like tag relatedness measure (1) that is based on the number of different contributors that used tags in pairs but also individually.

$$rel(T_1, T_2) = \frac{u(T_1, T_2) * log(u(T_1, T_2))}{u(T_2)} \qquad (1)$$

We define $rel(T_1, T_2)$ as the relatedness of tags T_1 and T_2; $u(T_1, T_2)$ is the number of users for which T_1 and T_2 co-occur in the same photo; uT_2 is the number of users having used tag T_2. The division by uT_2 penalizes tags used by many people in contexts different from that of the target tag T_1 and reinforces specific ones, which are often associated to T_1 and rarely associated to other tags. Relation 1 can be seen as a transposition of frequency-based models used in information retrieval to the social space. Instead of considering total term frequency, as it is done in systems where authorship data are unknown or neglected, we exploit community usage of a term in order to determine its pertinence in a given context.

Table 3: Top 10 related terms obtained for male (M) and female (F) users in our sample by applying relation 1.

Tag	F/M	Related Tags
Cactus	F	cactus flower, cacti, prickly pear, saguaro, opuntia, kaktus, echinopsis, cactaceae, prickly, succulent
	M	cactii, saguaro, cacti, cholla, spines, prickly pear, cactaceae, espinas, kaktus, succulent
Castle	F	castello, chateau, schloss, moat, castillo, burg, neuschwanstein, scotland, windsor, edinburghcastle
	M	castillo, burg, castello, chateau, schloss, castell, neuschwanstein, castelo, fotress, englishheritage
Food	F	cibo, salad, dessert, yummy, noodles, yum, zucchini, vegetarian, dinner
	M	comida, meal, cooking, eat, tasty, dinner, meat, fried, salad, steak

Table 4: Sample tags in cosine male-female similarity ranges.

Cos	[%]	Tag
$>$ 0.9	7	sky, blue, sunset, water, beach, tree, clouds, nature, church, trees
0.8−−0.9	29.6	red, green, flower, night, flowers, white, light, portrait, sea, yellow
0.7−−0.8	27.5	snow, cat, pink, macro, moon, lights, rose, paris, selfportrait, castle
0.6−−0.7	16.5	car, garden, bike, hat, painting, sepia, festival, guitar, cars, film
0.5−−0.6	9.3	friends, apple, camera, feet, photoshop, bokeh, underground, shopping, silver, cactus
$<$ 0.5	10.1	panorama, vintage, time, stars, fishing, flash, casa, airplane, studio, bag

In table 3, we present the top 10 related tags, for the 5 most frequent tags, computed using (1) over the sets of all female (F), or male (M) users. Related terms are often translations of the concept (castillo, schloss for castle), or hyponyms / instances (neuschwanstein, windsor, edinburgh castle for castle) or closely related terms when hyponyms / instances do not exist (dinner, tasty for food). The presence of many translations indicates that contributors often use several languages to tag one photo. Tag representation is closely related to automatic folksonomy extraction [9]. However the focus here is not on inducing hierarchical relations but on deriving gender-related sets of close terms. The results in table 3 indicate that related terms used by women (F) and by men (M) are different to some extent. We quantified the variation of gendered tag representations by computing the cosine distance between top 100 related tags extracted for female and male users (table 4).

The cosine male-female distance ranges with the highest concentrations are 0.8 thru 0.9 and 0.7 thru 0.8 (29.6%, respectively 27.5%). Only 7% of tag representations have a gender-neutral similarity score over 0.9 while the cosine dis-

Figure 1: Female results for food.

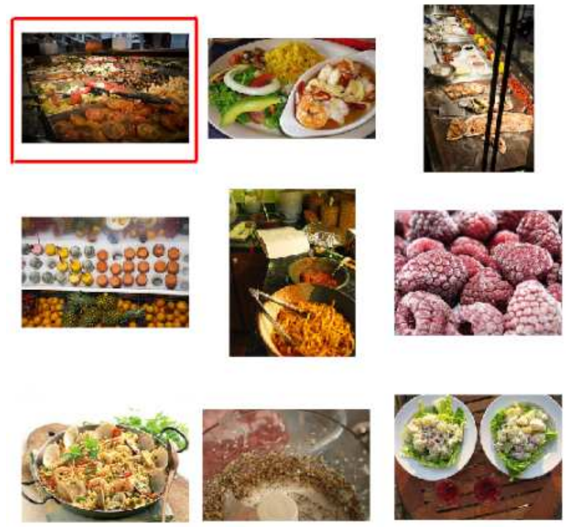

Figure 2: Male results for food.

tance is smaller than 0.5 in only 10.1% of the cases. We see that there are differences between women's and men's tag representations in a large majority of cases. Gender differences are not very salient for some general tags, e.g., sky (0.94), nature (0.92) or tree (0.94); and for some colors, e.g., blue (0.91) or yellow (0.88). Greater differences are obtained for terms such as fishing, stars or airplane. The results in table 4 indicate that artifact representations are not gender neutral, e.g., car (0.68), camera (0.53), airplane (0.47). Such comparisons of female and male tag representations help us understand differences between women's and men's world representation. These differences might be interesting in applicative contexts. For instance, the research presented here could help companies adapt product-related communication to female and male audiences.

6. GENDER SENSITIVE IMAGE SEARCH

The influence of users' gender on information retrieval processes has rarely been addressed [12]. We found no previous study that explores the adaptation of an image search engine to female/male audiences, so we created a gender sensitive image search prototype. To do this, we downloaded metadata for the top 20,000 Flickr images associated to the 150 most frequent tags in our 12,000 user sample.

Gender sensitive results were obtained using the following three steps heuristic:

1. For female users, and for male users, calculate the top 100 related terms for using formula 1 explained in the last section for each of the 150 tags.

2. For each tag T of the 150 tags, rerank the 20,000 Flickr images returned for this tag T in the following way. For a female user, consider the 100 closest related tags for T, let's call them F_1 to F_{100}. Now for each image them I_1 to I_{20000}, retain the first three tags from F_1 to F_{100} that appear in the tags for that image I. Let's call these tags F_i , F_j and F_k. Assign the following score to image I: $score(I) = \frac{1}{log(1+i)} + \frac{1}{log(1+j)} + \frac{1}{log(1+k)}$. Logarithmic values are used because linear weighting is too steep and would tend to over-represent a small number of top related terms. When ties appear (same 3 related terms), break the tie by using the Flickr rank of the images. As a result of this ranking, an image which is tagged with closely related terms will be ranked higher than another image which is tagged with more loosely related terms.

3. Diversify the reranked results using a list of encountered related terms. For each new image, test whether at least one of the 3 retained related terms is not in the list of already encountered terms. If so, add the image to the final results and add the new related term to the list of encountered terms. Diversification is necessary because simple query expansion favors the same strong tags. For instance, a query with cactus for male users will return a lot of saguaro images before diversification. After diversification, there is a greater chance that top results include one picture tagged with saguaro but also pictures of cholla or prickly pear. Strong diversification hurts precision and, as a compromise, we introduce a weak diversification constraint, with only one new related term sufficient in order to retain an image.

The procedure will rerank Flickr results according to preferential female tag representations, as well as for male tags. In figures 1 and 2, we present an example of gender sensitive results for food (a, female; b, male). Only one image (outlined in red) is common to female and male results. Female and male food representations tend to be reflected by results in figures 1 and 2. Expanded queries for female representations include (food, dessert, yummy, delicious) and (food, dessert, dinner, meal) while male expanded queries include (food, meal, dinner, seafood) and (food, cooking, meat, beef).

7. EVALUATION

We assess the influence of gender-adaptation on image search, as described above, using a two step procedure. A

Table 5: P@10 and standard deviations (StDev) for Flickr and gender sensitive image search obtained over a panel of 82 queries.

Method	P@10	StDev
Flickr	0.667	2.46
Female	0.849	1.77
Male	0.835	1.5

preliminary test compares precision at 10 results (P@10) comparing reranked results to original Flickr results in order to assess the overall quality of reranked results. Secondly, a user survey is performed to see if image search preferences are influenced by the user's gender.

7.1 Preliminary Evaluation

We start with the list of 150 most frequent tags which are used as queries because they are easily recognizable by almost any user and produce female and male reranked results using the algorithm presented in the preceding section. Next, we select the 82 image searches for which female/male reranking return at least five different results in the top 10 results. This selection is performed because we want to investigate those F/M results which are well separated. This sample includes terms as diverse as airplane, beer, hair, lighthouse, or Paris and is appropriate for testing the reranking method's generality. The test was blind, with top 10 image results for a query obtained with Flickr search and gender sensitive search randomly displayed on a same page. The evaluator (a man) had to choose relevant images for the query among those displayed and we report P@10 results based on his picks in table 5.

We performed unpaired t-tests to assess differences between the three sets of results and we found significant differences between the baseline and results reranked with female, respectively male related tags. The comparison of Flickr results (P@10 = 0.667; SD = 2.46) to female results (P@10 = 0.849; SD = 1.77) gives t(162) = -5.43 and p < 0.0001. The comparison of Flickr baseline and male results (P@10 = 0.835; SD = 1.5) gives t(162) = -5.29 and p < 0.0001. Results for female and male rerankings are not significantly different (t = 0.524 and p < 0.6) and, given that female and male result sets are well separated, this shows that our reranking technique is robust. Given that we report results based on a single participant's opinion, one could criticize the robustness of the quantitative evaluation presented here. However, we recall that precision is evaluated primarily in order to provide input data for gender sensitive adaptation testing.

7.2 Role of Gender

Out of 82 queries tested, P@10 scores are identical in 26 cases and we use these last queries to evaluate the influence of gender on the search process. Queries with identical precision are chosen because, since we next focus on assessing preference for female/male results, we need to minimize bias generated by differences of quality between female and male image sets. In this phase of the evaluation, results were gathered through an online survey which consisted of an introduction page, followed by 26 pages that presented top 10 female and male results on clearly separated but unlabeled columns, and a final page with explanations. The introduction page explained the task (but not the method) and also contained a short form which collected participants' gender, age and location. The position of F/M result sets on the pages was randomized in order to avoid the formation of patterns. Participants were asked to decide if they preferred one set of images to the other or if the two sets are equivalent (option was selected by default). From a list of 500 people contacted by e-mail, we received a total of 120 answers, out of which 78 were complete and could be analyzed. 48 participants were female and 30 male. The age distribution is: 60 from 21 to 30 years old, 16 from 31 to 40 years and two above or below; 50 participants from France, 18 from Romania and 10 from other European countries.

Table 6: Contingency table of preferences. F-res, M-res – preference for one gender's results; neutral – no preference.

Method	F-res	Neutral	M-res	Row total
Female	517	293	438	1248
Male	261	189	330	780
Col.total	778	482	768	2082
$gamma = 0.1379$; $tau - c = 0.088$				

In table 6, we present the contingency table of votes given by female and male participants. In a random setting, preferences would be equally distributed among the three columns. Since female and male results appear more frequently than neutral choices (38.4% and 37.9% vs. 23.7%), we notice that the adaptation of image results to gender is not indifferent to participants. To assess the overall association between gender and preferences for gender-adapted results, we calculate compute two measures of ordinal association (Goodman and Kruskal's $gamma$ and Kendall's $tau - c$). Both measures vary from -1.0 to 1.0, with: -1.0 a negative relationship; 0 statistical independence and 1.0 a positive relationship. The values obtained for $gamma$ (0.1379) and $tau-c$ (0.088) indicate that there is a weak positive association between image search preferences and user's gender. This translates the fact that, on average, each user of each gender chooses results of their own gender-reranked results more often than other gender's results. However, the influence of gender-reranking is highly variable from one participant to another and, in order to give an idea about individual variations, we present a histogram of individual choices in figure 3.

The results in figure 3 provide a detailed picture of the overall results presented in table 6. They show that around 2/3 of participants chose more results that correspond to their own gender's representation of the query and that 1/4 of participants choose more results reranked with the other gender's representation of the query. However, the preference for one type of results is not very well defined for around 2/3 of participants (dominance score between -3 and 3). Stronger preference (dominance score at least 4) is obtained for the other participants. One interesting observation is that female participants are more likely to manifest a clear inverse preference (5 participants) than male participants (0 participants).

The F/M results have equal P@10 scores and can be considered of similar quality. Another factor that might influence participants' choices is the gender predominance (Section 4) of queries themselves. To test this hypothesis, we computed a possible link between gender predominance and

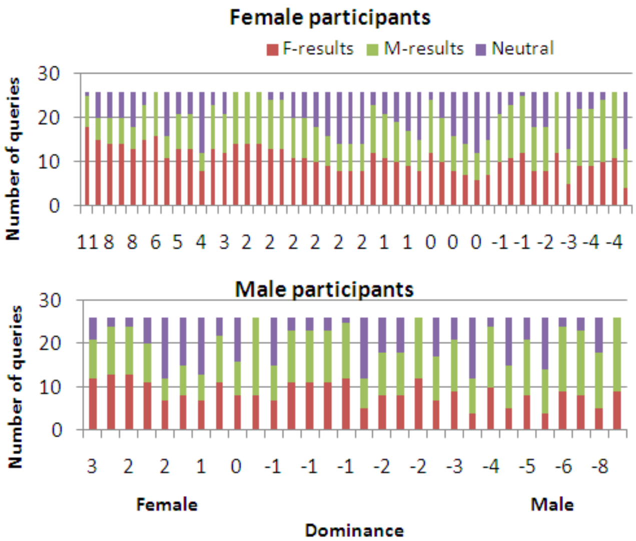

Figure 3: Distribution votes for female and male participants. The values on the X-axes stand for the strength of the preference for female or male adapted results.

preference for a type of representation but no significant association was obtained. With equal values of P@10 and insignificant influence of gender predominance, other factors must explain participants' tendency to choose gender reranked results. Other factors might be possible, but it seems that gender does have a predictable influence. Our experiment shows that around 1/3 of participants have a clear preference for their own gender's results. These users would be better served by image search results that are adapted according to their gender preference.

8. CONCLUSIONS

We analyze image tagging and search from a gender perspective. An analysis of a large user sample shows that there are significant gender determined differences of tag usage. Women and men photograph different things (to some extent) but also choose different tags for the same subjects, and analysis shows that, given a subject tag, supplementary tags added by females and males are often different.

Analysis of image tagging provides female and male tag models which can be exploited to create a gender sensitive image search prototype, providing results reranked through gender specific query expansion. A preliminary evaluation of 82 queries shows that substantial improvement is obtained for both female and male reranking of image search. Further tests show that around 2/3 of users prefer results generated using their own gender's tag representations. Moreover, around a third of the users express clear preference for their gender's representation. This indicates that image search is not a gender indifferent process and it could help improve the search experience at least for users with clear preferences.

With minor modifications, a procedure similar to the one described for image search adaptation can be used to obtain gender sensitive text search results. Another interesting development would be to evaluate the influence of other types of personal information, such as location or age as well as their combinations.

9. ACKNOWLEDGMENTS

This research is funded via the ANR Georama project (ANR-08-CORD-009).

10. REFERENCES

[1] S. Ahern, M. Davis, D. Eckles, S. King, and M. Naaman. Zonetag: Designing context-aware mobile media capture to increase participation. In *Proc. of PICS 2006 (UbiComp 2006)*, 2007.

[2] M. Ames and M. Naaman. Why we tag: motivations for annotation in mobile and online media. In *Proc. of SIGCHI 2007*, 2007.

[3] S. Argamon, M. Koppel, J. Fine, and A. R. Shimoni. Gender, genre, and writing style in formal written texts. *TEXT*, 23:321–346, 2003.

[4] T. Arni, P. Clough, M. Sanderson, and M. Grubinger. Overview of the imageclefphoto 2008 photographic retrieval task. In *CLEF 2008 Workshop Working Notes*, 2008.

[5] D. Crandall, L. Backstrom, D. Huttenlocher, and J. Kleinberg. Mapping the worldâĂŹs photos. In *Proc. of WWW 2009*, 2009.

[6] N. Garg and I. Weber. Personalized, interactive tag recommendation for flickr. In *Proc. of RecSys '08*, 2008.

[7] S. C. Herring and J. C. Paolillo. Gender and genre variation in weblogs. *Journal of Sociolinguistics*, 10(4):439–459, 2006.

[8] O. Nov, M. Naaman, and C. Ye. Motivational, structural and tenure factors that impact online community photo sharing. In *Proc. of ICWSM 2009*, 2009.

[9] A. Plangprasopchok and K. Lerman. Constructing folksonomies from user-specified relations on flickr. In *Proc. of WWW '09*, 2009.

[10] T. Rattenbury, N. Good, and M. Naaman. Towards automatic extraction of event and place semantics from flickr tags. In *Proc. of SIGIR '07*, 2007.

[11] E. Rich. Sex-related differences in colour vocabulary. *Language and Speech*, 20(4):404–409, 1977.

[12] M. Roy and M. Chi. Gender differences in patterns of searching the web. *Journal of Educational Computing Research*, 29:335–348, 2003.

[13] P. Schmitz. Inducing ontology from flickr tags. In *Proc. of Collaborative Web Tagging Workshop (WWW '06)*, 2006.

[14] B. Sigurbjörnsson and R. van Zwol. Flickr tag recommendation based on collective knowledge. In *Proc of WWW '08*, 2008.

[15] M. van Leeuwen, F. Bonchi, B. Sigurbjörnsson, and A. Siebes. Compressing tags to find interesting media groups. In *Proc. of CIKM '09*, 2009.

[16] R. H. van Leuken, L. Garcia, X. Olivares, and R. van Zwol. Visual diversification of image search results. In *Proc. of WWW '09*, 2009.

[17] L. Wu, X.-S. Hua, N. Yu, W.-Y. Ma, and S. Li. Flickr distance. In *Proceeding of ACM Multimedia '08*, 2008.