

Flickr Image Tagging: Patterns Made Visible

by Joan Beaudoin

Folksonomies

The development and subsequent popularity of image tagging at sites such as Flickr (www.flickr.org) has been a phenomenon to receive considerable attention over the course of the last few years. After spending more than a decade cataloging and providing access to images in an academic setting I, too, felt compelled to take a look at what this Web 2.0 image-sharing site had to offer. There were two interconnected ideas at play when I began thinking about performing a study of Flickr tagging. The first of these ideas had to do with looking for an underlying pattern for the image tags. It seemed likely that some commonalities would occur among what at first glance appeared to be the chaos of personally applied image tags. Finding these common patterns among the tags would clarify what types of information people typically associate with their images. The second idea concerned the effectiveness of image tagging. If patterns were discovered among the image tags, I believed these could be used to alleviate some of the problems associated with tagging. So early in 2006 with these two ideas fresh in mind, I carried out a small study of the image tags used at Flickr.com.

To conduct the study I gathered the top 10 image tags of 14 randomly chosen Flickr users and downloaded them through the site's open APIs. In order to discern if there were patterns to be found in the application of the image tags, I applied conceptual labels to each of the 140 image tags. This labeling process was iterative, and after several passes through the entire set of image tags a model consisting of 18 categories (Table 1) emerged. During this process it became apparent that image tags could have multiple meanings and as a result, I allowed some to be assigned to several categories. Thus, a tag such as *cross* could simultaneously be considered a verb, a thing, an emotion and an adjective.

Joan Beaudoin is a doctoral student at Drexel University. She can be reached by email at jeg56@drexel.edu.

Model Evaluation

In order to evaluate the usefulness of the model to represent the various image tags, I gave the files of the extracted Flickr image tags and the categories with definitions to four people, who then categorized the image tags in a way that made most sense to them. Their categorizations of the image tags were combined with mine and the occurrences were then tallied to determine the overall category agreement (Table 2) and patterns of category usage (Table 3).

TABLE 1. Category model for image tags found in Flickr.

Category	Definition	Examples
Adjective	All adjectives	cold, wet, bright
Compound	Terms with two or more words combined	Newyorkcity, Mydog
Emotion	Identification of emotional state	happy, depressed
Event	Pertaining to holidays, happenings or news occurrences	wedding, Easter, assassination
Humor	Terms used for humorous reasons	lthinkbobbyisgreat
Language	Terms in any language beyond English	eau, gefühle, madrina
Living Thing	Living, non-human creatures and plants	bird, rose, tree, dog
Number	Terms composed of numbers	64325, 1+111, 2000
Person	Named (common and proper) individuals and groups	baby, Elvis Costello, Girl Scouts, woman
Photographic	Terms relating to imaging/photographic devices and, or processes	Cannon, SLR, l100
Place-general	Places identified with their common names	beach, field, bedroom
Place-name	Places identified with their proper names	Amsterdam, Seoul
Poetic	Terms that are poetic in nature	heavenly mirage, daydream
Rating	Terms which evaluate images	topten, tag1, taggedout.
Thing	Non-living objects	house, car, rock, water
Time	Terms with chronological meaning	June, 2006, night
Unknown	Unidentifiable terms	Sha78, Pp73
Verb	All verbs	running, look, crying

BEAUDOIN, continued

TABLE 2. Percentages of agreement in category assignment.

Category	Percentage
Place-name	85.76
Person	75.24
Event	72.86
Verb	68.00
Emotion	60.00
Compound	58.60
Adjective	58.33
Place-general	55.56
Thing	53.68
Unknown	50.59
Living Thing	46.67
Photographic	41.00
Humor	40.00
Number	33.33
Language	31.11
Rating	30.77
Poetic	30.00

The overall category agreement of the model suggests that the model is modestly effective in describing the concepts of the tags as assigned by Flickr users. Five people agreed on 11 of the 18 categories more than 50% of the time. Several of the lower performing categories (photographic, language, rating) seem to have been the result of a lack of specialized knowledge on the part of several of the participants. For instance, language tags (Weg, tio) were sometimes placed in the category *Unknown*. One category, *number*, seems to have had an unclear meaning for several people, and several tended to use the category *time* for tags consisting of digits, which could account for the lowered agreement in this case. The categories of *humor* and *poetic* also performed

poorly, but the reason for this situation is unclear. It seems likely that these categories of tags were the most open to individual interpretation, however. The model's performance could be improved through modification of those categories with low percentages of agreement across participants and by providing some further instruction and examples of typical tags to be found in each of these categories. An additional means of testing the model would be to instruct the participants to choose a single, best category for each tag.

Tag Categories – Frequency of Use

Looking at the percentages of tag category usage it is clear there are preferences for tagging among the Flickr users. The most frequently used categories of tags by this group of Flickr users were named geographical locations (New York City, China). Compound tags (white dish, rock star glasses) were the second most frequently occurring category of tags among these Flickr users. However, unlike the other highly occurring tag categories

in the model, which are descriptive in nature, the *compound* category is applied for reasons of form alone. With these tags the user applies a tag that combines two or more terms. The high percentage of *compound* tags suggests that Flickr users find single word tags to be inadequate for describing their images. The next most used categories were inanimate things (water, bottle) and people (Debbie, woman).

After these high performing categories of image tags follows a group that saw only modest usage. These tags each accounted for approximately 1-5% of the total tag use by the Flickr users. Even though their usage appears to be limited, their overall importance should not be underestimated. For example, the category *event* (party, wedding) appears in each of the users' top 10 most frequently employed tags. This usage illustrates the value people place on identifying special occasions within their images. The relatively modest use of this category among the tags explored here I

believe speaks more to the rarity of these events than it does to a lack of interest in tagging images with these terms. One only needs to look at the tag cloud of Flickr's most popular tags to see the frequency with which *event* tags are used. While the list of most popular tags reflects those terms that are most frequently applied by the entire Flickr community it is suggestive of how common these terms are as image tags.

At the lowest end of the spectrum of tag usage, all occurring at below 1%, are the categories *humor*, *poetic*, *number* and *emotion*. These categories plainly see little usage among the top 10 tags of these Flickr users. However, it needs to be mentioned here that the full range of tags being assigned by users was not investigated. It may be that these types of tags are being employed at a

TABLE 3. Percentages of category usage for image tags found in Flickr.

Category	Percentage
Place-name	28.21
Compound	14.05
Thing	11.37
Person	8.81
Event	5.69
Unknown	4.79
Photographic	4.57
Time	4.46
Adjective	3.90
Verb	3.79
Place-general	2.79
Rating	2.23
Language	1.56
Living Thing	1.56
Humor	0.67
Poetic	0.67
Number	0.56
Emotion	0.33

higher frequency but that they are not well represented in the users' top 10 tag lists. Whatever the cause of this situation, the phenomenon is unclear and warrants further research.

Problems of Meaning

A further issue needing additional study is the difficulty of discerning the meaning of some image tags. Even when the participants could apply ambiguous tags to several categories, they could not characterize a high percentage (nearly 5%) of them. In some cases these *unknown* tags appear to have a personal meaning for the Flickr users who applied them (for example, "tg78") and so the participants performing the categorization could not readily understand them. In other cases the *unknown* tags illustrate how important contextual knowledge is to the categorization of the tags. This factor was most clearly seen in the tags relating to Flickr groups and photographic devices, which were unfamiliar to several of the participants. Foreign languages and unknown geographic locations posed recognition difficulties for some participants as well, so a number of these tags also found their way into the *unknown* category.

Individual Differences

The tag categories discussed above give an overview of the way Flickr users tag their images as a group, but their application by single users showed some degree of variation. For example, several users had a single instance of *place-name* in their top 10 tag lists, while others had top 10 tag lists consisting nearly exclusively of this category of tag. The categories *time* and *photographic* also saw highly individualized usage. While tags with photographic or imaging concepts behind them were used by a majority of Flickr users at least once in their top 10 tags, for one individual this type of tag accounted for half of the top ten tags applied. The frequency of this tag suggests that most Flickr users are interested in recording the processes and devices used in the creation of their images, but for some users it was a far more important form of information. The use of *time* showed the most variation among the Flickr users that were studied. Eleven of the 14 users

did not have a single instance of *time* among their top 10 tags, but three users found it to be a useful categorization, and for one user it accounted for six of the 10 top tags. The limited use of *time* by Flickr users is an interesting discovery since it is a frequently used organizational principle associated with the management of images. This situation is possibly due to the fact that image files commonly receive a date stamp when they are created. Therefore, some Flickr users would see this information as redundant.

Flickr Developments

Interestingly, over the course of the year that has passed since I performed the tagging investigation Flickr has implemented several features that support many of the categories of information this study found to be important to users. When files are uploaded to the site the date stamp (partially representing the information stored in the category *time*) now automatically displays alongside the image. In addition, the device used (partially representing the information stored in the category *photographic*) to create the image is also automatically recorded as the file is uploaded to the site. *Place-name*, the category that saw the highest frequency of use, is now partially accommodated through the use of geotagging, and as a result a named geographic location is now displayed alongside the image. To take advantage of this feature users place their images on a map manually, or they use an application written for camera phones which records the geographic coordinates of the location of the cell phone's tower for photographs uploaded to Flickr. Photographs that have been geotagged using either of these methods can be viewed on a map launched through Flickr that can also display the locations for other nearby photographs that have been uploaded to the site. Three other categories that were highly used, *thing*, *person* and *event*, haven't seen similar support from Flickr. However, the tag category *compound* has seen the development of a technique that now allows for multipart tags to display as separate words in the user's tag list. By inserting double quotation marks around the entire tag, users can cause compound tags to display as they are entered (for example, Father Time instead of fathertime). Each of these improvements to the Flickr site offers additional

means by which to record information about images and in several cases these require little or no effort on the part of the user. Not all Flickr users apply tags to their images, and so in these cases this automatically recorded information may be the only means of access beyond visual browsing.

Implications for Users

The model revealed several important aspects about the tagging behaviors within the Flickr community. The identification of the types of information being recorded for images is obviously a useful step in helping to develop more effective methods of tagging. Flickr has been hard at work in this regard, as was noted above. In addition to streamlining the tagging process, the automatic ingest of information that happens with image uploads to Flickr deepens the information pool and removes human error. Further developments could be implemented to decrease the effort involved in tagging. The most basic of these is reducing the cognitive load associated with the task. A basic schema developed from the model could be employed to prompt individuals to enter tags rather than trying to choose the “right” words to represent the image. Associating tags with conceptual headings would also prove useful for clarifying meaning (that is, does the tag *cross* signify the verb, the thing or the adjective). Tags using conceptual headings could have thesauri associated with them to facilitate the choice of additional tags. This area is one I think would see heavy use by Flickr taggers. One of the most interesting aspects of tagging witnessed at Flickr is the care some users take in applying as many tags as possible for a single concept. For example, an image may be tagged as cat, kitten, feline, or felis silvestris catus in addition to adding tags in multiple languages: chat, gato, gatto, Katze. In addition, each of these tags is often entered again in its plural form. An automated tool to accommodate the plural or singular form

of a tag would obviously be useful in these cases. Combining tagging with a basic model enhanced with tools like those discussed above would serve to strengthen the natural language of the users’ efforts and help to increase the retrieval of relevant images.

Implications for IS Professionals

The popularity of image tagging is a testament to the effort people are willing to expend in describing their images. As information professionals we need to develop new methods and techniques to assist people in tagging and retrieving their ever-growing body of visual materials. This information will in turn inform our own image indexing practices. With a better sense of how individuals are tagging their personal images it would behoove us to offer a similar kind of information for the visual materials we provide access to in our own collections. If we are not recording similarly detailed, descriptive information concerning the places, people, things and events in our image collections we are probably not reaching the broad audience we all hope to serve. Additionally, this study should clarify for information professionals just how highly personalized image information can be among individuals. Although a general pattern of image tagging was discerned across the Flickr users, information an individual was interested in might focus on a single aspect. This discovery is indicative of just how similar visual information research needs are to text-based investigations. Visual information is as richly complex as text-based materials, but at present text is more accessible. Concurrently with the technological developments over the past decade, there has been considerable progress in image retrieval, but a great deal remains to be done. One way that we can make these improvements is by continuing to look closely at how individuals categorize their own images and to assist them in their efforts. ■