Do Tags Work?
*Cathy Marshall*

Tag! You’re it!

It seems that everywhere I go on the Web these days is tagged.

When I login to Flickr, the first photo I see—an artistic effort that looks to be a study of a sturdy dandelion growing from a crack in the curb—is tagged *flowers*, *spring*, *scenery*, *lapsana*, *great*, *fotolog*, *ThinkFlickrThink*, *SuperShot*, and *1mill*. I like the photo, although it calls to mind dozens of inspirational posters, the province of the overearnest, the literal, and the hapless striver. *Flowers*? I can see only one. A *lapsana*. Who knew? Is it *scenery*? Well, it’s not a dog, although I think of scenery as grander in scale and as less ambiguous in its beauty than a lone dandelion. Is it *great*? According to his comment, Blue Cockatoo thinks so. How thick is the stem? Oh, I’d say it’s about *1mill*.

Heckuva job, tags!

Even *The New York Times*, bastion of old media order and civility, is more than willing to present you with a tag cloud view of the news. Is that a deluge of *bobby fischer* news I see on the horizon in the wake of his untimely demise? Or perhaps it’s more like a fine mist of *bobby fischer* rumors. There must be a review of a *japanese cell phone novel* about *modern love* among the *gay ufos* in the news today too. If we believe the cloud, *hillary* Clinton is becoming more and more like Madonna and Cher, so recognizable that she no longer needs a surname.

And then I think to myself, “*suzanne pleshette*? What’s Suzanne Pleshette doing there? Did she die, or has she entered the presidential race?”

With tags I’m right on top of the news.

The promise of all these tags—the gathering clouds on the horizon, a harbinger of what can only be intellectual global warming—is nothing short of a full-fledged folksonomy. Large segments of lawless Internet will become manageable and accessible in a way that is out of the reach of underfunded libraries and other institutions charged with
information stewardship. By harnessing the wisdom of crowds, this roiling sea of knowledge will be calmed; its power will light a billion tiny “aha!” light bulbs over a billion tiny avatar heads.

But are the tags that people create really an effective way of describing information so that it can be found and managed, folded and put in the right drawer?

I can’t deny the populist appeal of tags. People adore tags. In principle, at least.

David Weinberger is perhaps the most articulate and outspoken proponent of the pro-tagging point of view. I went to a talk he gave at the University of North Carolina in the wake of the publication of his wildly successful new book, *Everything Is Miscellaneous*. He was awfully convincing. I’ve looked through enough folders with “misc” on the tab to understand how often things fall through the epistemological cracks. There’s no question that people would be better off assigning tags to classify their own stuff. And I don’t doubt that information wants to flock.

Things often look different when you put them next to one another; witness the fashion statement you can make by dressing in the dark.

But I believe Mr. Weinberger actually used the embarrassing words “stick it to the Man” when he was extolling the virtues and power of social tagging. It reminded me of nothing more (and nothing less) than the vivid funeral scene in that so-bad-it’s-good Peter Fonda movie *The Wild Angels* (1966). Tags, in this construction, are the surest antidote to decades of oppression by librarians and other elitist thugs of the information establishment.

*Stick it to the Man.*

I haven’t heard anyone say “Stick it to the Man” for a long time. If you could’ve seen me in the auditorium listening to David Weinberger, you surely would’ve noticed me squirming in my seat. And you would’ve seen my telltale thought bubble that recalled a version of Heavenly Blues’ (Fonda’s) eulogy of The Loser (played by Bruce Dern):

“We want to be free! We want to be free to do what we want to do! We want to be free to tag. And we want to be free to mash up our web sites without being hassled by The Man.”

You see where I’m going with this. I’m just uncomfortable with the amount of power Weinberger has ascribed to social tagging. Has he looked at real tags?

For that matter, have I?
I’m convinced that tags provide us with a fine way to organize our own stuff. After all, I was a member of the Hypertext community before stuff-organizing was fashionable, back when faceted classification was an obscure idea attributed to an Indian librarian named S. R. Ranganathan. Even without facets, you don’t have to look very hard to see that people seem to function pretty well in a world full of things that they’ve organized all by themselves—grocery lists they’ve written on the back of envelopes and to-lists based strictly on the satisfaction they get from crossing off things—without leaning on the tricks espoused by Lifehacking gurus like Danny O’Brien and Merlin Mann.

But I do need to be persuaded that tags are of use to strangers. I’m no Blanche Dubois of the data glut.

How might I find out whether people can tag worth beans?  

Here’s what I did: I collected 322 public images from Flickr—photos of a remarkable mosaic inset into the floor of a famous Milanese Galleria—and looked at their tags.

Of course you could say, even inadequate tags are better than nothing at all. But I wondered if perhaps my fellow Flickr contributors are more adept at other sorts of description than they are at tagging. Are they better at giving pictures titles—for example, that dandelion photo is aptly titled “Lapsana apogonoides”—or writing narrative descriptions of them than they are at tagging them? Is all the effort that goes into social tagging paying off? So besides grabbing the images and their tags, I also gathered the titles the aspiring photographers assigned to their photos and harvested the brief narratives they wrote about them.

I became a metadata gleaner. I gathered and harvested; I tweaked and fiddled; I put all the data in a big spreadsheet and used pivot tables; I counted what was countable; and I laid out things side-by-side to compare what was comparable.

What I found surprised even me.

I’ve been using Flickr as a source of stock photography for a while now, so even at the outset of this experiment, I knew two important things that not everybody realizes yet: (1) There’s no longer any point in taking snapshots of places and things; someone’s already taken exactly the picture you want to take. The lighting’s just like it would be in your photo. The subject is just as out of focus and just as poorly framed as if you’d have taken the picture yourself. Even the people are the same: their friends look just like your friends. (2) There’s not just one photo like the one you would’ve taken yourself had you

---

1 I’ve already had my brief foray into the much vaunted ESP game; I was shocked how little value those tags had. That’s a pathological case though: people don’t care about what they’re tagging, and that’s sure to have an effect on the absolute worth of their tags.

2 I did not, however, harvest the comments that accrue; that we’ll leave for another study.
remembered to take your camera with you and charge its batteries. There are many photos like the one you would’ve taken yourself. Many, many photos. Not one. Not ten. More. A lot more.

Here I need to digress to tell you something important about me: I’m compulsive. Very compulsive. When I hear David Sedaris talk about his many tics and obsessions like licking light switches (when he was a child) and touching the top of peoples' heads (still), I register full identification. Once I start some kind of fussy project, especially one that involves collecting and counting, I can’t stop. You need to be very fussy indeed to collect 322 photos of a particular scene from Flickr, especially if you need to scrape the screen to get the ones that have been copyrighted. “Hasn’t anyone ever heard of fair use?” I grumble as I type Alt-Prt Sc Cntl-V Cntl-S to snag a photo whose owner thoughtlessly protected it against people like me.

I don’t remember how I first came upon this bit of Milanese tourist lore in Flickr—I think I was looking for a photo of those adorable faux bull testicles one hangs from the bumper of one’s Chevy Tundra—but once I’d seen a version of this photo in Milan and read the story that went with it, I knew it was just the example I needed to investigate the lure and efficacy of tags.

To quote antistar in Virtual Tourist, who wrote this comment under the heading of “Galleria Vittorio Emanuele: Turin Bull”:

Right underneath the glass dome in the Galleria Vittorio Emanuele is the emblem of Turin, a bull, on the mosaic floor. It is meant to bring you good luck if you spin around on your right heel on the bull's most treasured possession. I kid you not, there was a crowd of people queuing up to crush the testacles [sic] of the bull of Turin, in order to bring themselves good luck. I don't know if this is a tradition imported from Turin, or if the Milanese are showing a great disrespect to their neighbours, but it was a fun to watch people grinding their heels on the poor bull's private parts.

It didn’t just seem like a good example; it seemed like the perfect example. Sufficiently remarkable that the mosaic itself would command the attention of scads of travelers. And it has a stock bit of folklore to go with it. What’s more, there’s a ritual associated with your visit—you spin around three times, on your heel, counterclockwise, and possibly make a wish. Would you wait for someone to spin on the mosaic so you’d have

---

3 Type “bull balls” into Flickr’s search box to see for yourself: you’ll find dozens of photos of the Milanese bull intertwined with New World photos of the Wall Street bull. I decided the Milanese mosaic was a better subject for this investigation.
a person in the picture? Would you annotate the growing cavity in the ground? Finally, any description would be a nuanced dance around taboo language: would you say “Bull balls”? Would you be serious and anatomical: “bull testicles”? Would you avoid the issue altogether and just use ellipses or cutesy allusions (as many did)?

For several weeks, I fished in Flickr’s photo stream. Compulsively. Obsessively. Fished for the bull. Foregoing my blog, my job, and even the meowing cat, who clamored for my attention by lacerating my feet with his razor-sharp front claws.

My patience rewarded me. I initially found 322 recognizably similar photographs of the Turin Bull mosaic in Milan’s Galleria Vittorio Emanuele. Several days later, I found 4 more, added in just that interval. I’m sure there are even more now, today.

I decided to close the collection at the original 322, since it’s obvious that I’m looking at a moving target.

Here’s what they look like. Some are just pictures of the mosaic. They look something like this:

Others focus on a person—sometimes a cute girl, sometimes a nebbishy guy—enacting the tale’s ritual, spinning for good luck.

---

4 In fact, I just redid one of my more successful queries—Milan bull—and ranked the results by date. There have been 70 new photos of the mosaic added over the last year, since I gathered the data for the study. Eight of them were added in the first two weeks. If you look under lucky bull, there are yet four more. See what I mean? You could make yourself crazy. Especially if you’re not fond of pit bulls.
Still others zero in on the action, the heel in the hole, poised to spin.

In spite of these minor variations (which I duly noted in my data analysis), I think of the pictures as all being quite similar: if I were looking for stock photos to illustrate a blog post, I’d think any of these variations to be adequate.

You might well question how I gathered this dataset. I gathered it patiently. I cast a broad net; I looked through thousands and thousands of photos. I used two and three word queries, combining words like Milan, bull, balls, spin, luck, mosaic, Italy, turn, travel, trip, and more. I ventured into Italian too, going out on a linguistic limb with words like toro, fortuna, volver, palle, Milano, Italia. Oh, and I queried more specifically for various combinations of Galleria Vittorio Emanuele II (say, “milano galleria” or “Vittorio toro” or “Galleria Vittorio”).

As you can see, I was neither overly precise nor so persistent as to look through every photo that claimed to be about Italy. I was thorough within limits. My records showed that I mulled over the results from 36 separate queries, although some of the broader ones were abandoned after I’d gone through the first 2500 photos or so. Many times I’d encounter photos I’d already recorded in my hand-constructed log. In fact, that’s how I decided to stop—when queries stopped turning up any additional examples. I grew

For those of you who want to know more precisely what I did, I used four types of queries to bring back sets of photos from Flickr, 23 queries that were in English; 2 queries that were place-only; 9 queries in broken Italian; and 2 queries that were mixed English-Italian. The English-language queries were: Milan wish, Milan bull, milan bollocks, milan spinning, milan turning, Milan luck, Milan mosaic, Milan trip, Milan lucky bull, Milan heel, Vittorio bull, galleria bull, galleria mosaic, spin 3 times, spin three times, spin luck, spin heel, heel turn mosaic, lucky bull, lucky bull Italy, bull balls, bull turin, bull mosaic. The three queries that were place only were: Milano galleria, Galleria Vittorio Emanuele. My unfortunate use of Italian resulted in these queries: toro Milano, toro palle, vittorio toro, toro volver, toro galleria, toreto, fortuna Milano, fortuna toro, porta fortuna. The mixed queries were: toro milan, fortuna milan. If the query results were high recall, low precision, I looked through the first couple thousand to determine whether the query was turning up anything new. For example, “Milan trip” turns up almost 5500 photos. I looked through the first 1500 and found 2 photos of the bull. In most cases, the sparse positive results were familiar ones, photos I had already gathered. If there were fewer than 1,000 results returned, I looked through all of them. If there were more, I seldom ventured beyond result 2,500. A high-precision query would be one like Milan spin, where 69/99 of the results are pictures of the bull mosaic.
familiar with specific photos: the low-resolution one from a camera phone; the one with the sheepish-looking man in the hat; the summer ones where Midwestern subjects are clad in horrific tourist regalia, tank tops, shorts, and the like as they spin; the winter ones where subjects are bundled up against the surprising cold, their mufflers flaring out as they twirl around on the bull.

It was not wholly scientific, this gathering, but my method went far beyond anecdotal. If image recognition worked better, I could’ve gathered all of the matching photos. If I were perfectly patient and had all the time in the world (or threw the search over the wall to the Amazon’s Mechanical Turk), I could’ve ensured complete coverage at a given point in time.

But this example is intended to sway you, not to determine empirically how many public photos of the darned bull exist in the Flickr database. As I looked harder and harder, any sort of metadata describing the photo became thin on the ground—I could’ve added more examples that had no tags, no title, and no narrative description, but I don’t think it would have either bolstered my argument very much nor detracted from it. If anything, it would just demonstrate that there are lots of photos with neither tags nor descriptive narrative, titled exactly as they came from the camera.

Thus I came to have 322 examples of the photo, coded, and organized in a spreadsheet. As I said earlier, I gathered not only the photos, but also the titles, the narrative descriptions, and of course, the tags. This felt to me to be enough data to conduct a perfectly credible preliminary investigation—it was enough to convince me.

So first let’s look at what’s there to analyze. How broad of a net did I cast? You’ll know the answer to this question by how many bull photos I uncovered that had minimal metadata. First of all, as you’d guess, there were no photos that were missing all three sources of description. That makes sense. How would I have found them? It would’ve meant combing through all the untraversed reaches of Flickr or, at the very least, scrounging through all of the Italian photo sets. Yet—to demonstrate the sincerity of my search—there were four photos with no titles, and 25 more that had only the names that were assigned by the digital camera software (e.g. DSCF3091.jpg). Some of these had a paucity of other description as well. For example, I unearthed a photo of the mosaic with no title, no narrative description, and eight somewhat generic tags: milano, milan, italy, italia, mosaic, galleria, italie, Natalie. Another photo without title or description was tagged: galleria, mosaic, milano, milan, shoes, converse, WS. I searched with due diligence.

To give you a further sense of the minimal end of the spectrum, I’ve counted the number of photos without: without tags, without titles, without description. I’ve broken these counts further by whether they’re of the mosaic, of a person standing on the mosaic, or of some anonymous leg in place to begin spinning:
Thus I’ve selected photos that are findable using metadata. This approach makes sense: my enquiry is about social tagging, not about labeling one’s own photoset for one’s own use. You might be able to remember that you went to Italy in August, 2005 and that you took a picture of the bull mosaic, but I’d never be able to recover that picture. So I’ve cast a broad net to select 322 findable photos.

The differences are sufficiently crisp that I feel it’s safe to say that people are not apt to deposit a public photo without assigning a title to it. Only 1% of the total photos I collected are unnamed. Next come tags: somewhere between one-quarter and one-third of the photos are untagged. This seems to indicate that people do feel that tags help—that they’re more vital for retrieval than the narrative, which is missing from roughly 40% of these photos. And this photo subject represents a best-case scenario for eliciting narrative: there’s folklore behind the bull mosaic; there’s something to explain, especially in the case of action or mosaic photographs, which don’t have a human subject to motivate them.

I had wondered when I started categorizing these photos whether people treated photos differently if they had a person in them. So far, the difference is negligible. If someone’s going to assign metadata to their pictures, they aren’t going to worry whether they’re explaining a person, place, or thing (roughly, my three photo categories).

Let me tell you a little more about the photo metadata I’ve collected. For the time being, let’s ignore any missing metadata—absent titles, missing narrative, tags unassigned—and see how much description people contribute when they bother to contribute anything. In other words, let’s see how many tags people contribute when they bother to make up tags, how much they write when they take the trouble to produce a narrative, and how long their titles are. Table 2 sums up the descriptive metadata; the numbers (words per narrative, words per title, and tags per photo) are means and the standard deviations are in parentheses.
But what do these numbers mean? Here’s a sample 22 word narrative from the dataset:

“So in Milan they say that if you spin on your heel three times on the Bull’s balls, you get good luck.”

This is actually quite typical of the narratives that accompany the picture, give or take a name or two. A 34 word narrative is not wildly different:

“At the center of Galleria Vittorio Emanuele II, it is said that you will have good luck if you step on the Taurus' testicles (not a real taurus, just a mosaic) and turn twice!”

A 51 word version?

“Planting your heel and twisting with a flourish on the ‘private parts’ of the mosaic picture of the bull in Galleria Vittorio Emanuele, Milano, is a tradition for Italians and tourists alike. The bull offers good luck. A nun just did it, the other can't miss doing it but is hesitant.”

You get the idea.

Titles? Titles too have their regularities. Two typical four word titles are “The Bull in Milan” and “Galleria Vittorio Emanuele II”. Seven words gives you room to be a little more expansive: “Rotate on the Bull’s Balls for luck” or “Bull Mosaic in Galleria Vittorio Emanuele II”. Sure, there are variations. But what surprised me more is how similar they all are. Here are all 37 titles beginning with the letter S. Some of the repeated titles are the apparent result of copy-pastes, but many are just similar.

<table>
<thead>
<tr>
<th>Seeking for luck</th>
<th>Spinning on the bull’s balls</th>
</tr>
</thead>
<tbody>
<tr>
<td>Si quiere volver a Milán: písele las güevas al Toro!!!!</td>
<td>Spinning on the bulls balls...its supposed to give me luck</td>
</tr>
<tr>
<td>Simon on the bull’s testicle</td>
<td>Spinning on the toretta</td>
</tr>
<tr>
<td>Simona spinning</td>
<td>Spinning your heel on the bull's testicles is apparently good luck in Milan</td>
</tr>
<tr>
<td>Some lucky bull...</td>
<td>Squashing the bull’s balls</td>
</tr>
<tr>
<td>Spin</td>
<td>Squashing the bull’s balls</td>
</tr>
<tr>
<td>Spin 3 times for luck</td>
<td>Squashing the bull’s balls</td>
</tr>
<tr>
<td>Spin for luck</td>
<td>Squashing the bull’s balls</td>
</tr>
<tr>
<td>Spin for luck</td>
<td>Squashing the bull’s balls</td>
</tr>
<tr>
<td>Spin Three Times And Good Luck You Will Find</td>
<td>Step and Spin!</td>
</tr>
<tr>
<td>Spinning</td>
<td>Step on the Bull’s Balls for luck</td>
</tr>
<tr>
<td>Spinning</td>
<td>Step with your heel on the bull’s balls, turn around and do the chicken dance for GOOD LUCK!</td>
</tr>
<tr>
<td>Spinning after a wish</td>
<td>Step with your heel on the bull’s balls, turn around and do the chicken dance for GOOD LUCK! (1)</td>
</tr>
<tr>
<td>Spinning for luck</td>
<td>Stepping for good luck</td>
</tr>
<tr>
<td>Spinning on bull testicles in Milan</td>
<td>Stepping on the Balls of the Bull</td>
</tr>
<tr>
<td>Spinning on Taurus the Bull</td>
<td>Stepping on the bull’s balls</td>
</tr>
<tr>
<td>Spinning on the Bull for Good Luck in Milan</td>
<td>Superstition</td>
</tr>
<tr>
<td>Spinning on the bull in the Galleria Vittorio Emanuele II</td>
<td>Susie Stepping on the Bull - Galleria Interior - Milan</td>
</tr>
</tbody>
</table>

Table 3. Tags beginning with S
Tags—which are by definition classifiers and therefore intended to be regular—are where things start getting wild and wooly. There are enormous variations in tagging strategies: Do I tell you where the photo was taken, what’s in it, where I’m going to go next on my vacation, or what I’m doing in Milan? The crowd that believes wholeheartedly in tags will tell you that a tag is relative to the person doing the assignment, so this variation isn’t surprising.

In fact, if we relegate tags to a minor role in personal information management, individual variation isn’t a bad thing. Perhaps the tags are simply well-tailored to their use. While {Arriving Home, Dinner With Lia, Milan or Bust} is an awful tag set if one is attempting to use Flickr as a database of stock photography, it may be just right if Bill is trying to find those photos he took on his vacation to Milan, the ones from the day he had dinner with Lia. But are the bulk of actual tags this personalized and quirky?

To set up this discussion, it’s important to realize that short tag sets (5 tags or fewer) are much more common than long tag sets (10 tags or more). Recalling that 97 of the photos have no tags, the short version of the story is that 150 of the photos have 1-5 tags; 54 of the photos have 6-9 tags; and 21 have 10 or more tags. You can picture what this drop off looks like:

Now, let’s look at some examples of two word tag sets: {milan, italy}; {milano, me}; and {Nat, Nathalie}. In fact, the most common two word tag set turns out to be: {milan,

---

6 This is an actual 3 tag set used to describe a photo of our canonical bull mosaic.
Italy); 20 out of 48 are exactly that. What about four word tag sets? One common one is {Milan, Milano, Italy, Italia}. Others illustrate the heuristic that tags are indeed relative to the tagger: {friends, vacation, milan, italy}.

Nonetheless, all of these tags are startlingly generic (that is, they are the same across many different photographers) and likely to be less informative than the other sources of metadata (including their grouping into photosets, which can be used to partition one’s own photos into a group taken in Milan). The four word title “Bull Mosaic in Milan” is a much better classifier (from the standpoint of retrieval) than the four element tag set {Milan, Milano, Italy, Italia}, which produces results so low on the precision-recall curve that the desired bull mosaic is lost in a sea of irrelevant photos of the countryside and the latest fashions.

What happens when tag sets become longer? After all, the average narrative is over 20 words long. Maybe a longer tag set will yield a more precise description. Unfortunately, it is readily evident from the data that this is not the case. First, consider that there are only 21 photos with 10 or more tags assigned to them. That’s under 7%. It’s more usual for photos with tags to only have one or two. Second—and this is unlikely to surprise you—photos with a larger number of tags conform roughly to the profile we’ve already developed for typical metadata. They may even be slightly better described and more fully titled than average. In other words, more tags don’t mean less narrative. Quite to the contrary. The people who apply more tags also write richer descriptions and are no less likely to make up titles.

In fact, let’s pull out the 21 cases where the photos have 10 or more tags. Seven of these 21 photos (33%) do not have narrative description, a figure slightly lower than the overall average of 39%, but not shockingly different. Certainly given the size of the sample, it’s reasonably consistent. Likewise, only one lacks a title. What about the length of the narrative when it is present? 36 words long—half again as long as the ones with fewer tags. And what about the titles? About 6 words long—again, 50% longer than the ones with fewer tags.

Perhaps the people who assign lots of tags are just a wordy lot.

But it’s quality, not quantity, that matters here, right? What kind of tags are in these lengthy tag sets? Are they comparable to good narratives?

At least a few are. Here’s a twelve tag set: {Italy, Italia, milan, milano, bull, good, luck, galleria, vittorio, emanuele, II, mall}. Not bad. But nonetheless, keep it in mind. We’ll come back to it. There’s something very important missing.

---

7 One outlier has 30 tags, but this is at a pathological remove from the others. That’s 12 more tags than its nearest neighbor with 18 tags.
Most of the large tag sets are not as informative as the example I just cited. It’s as if the photographer is struggling to come up with meaningful descriptors. Here’s a twelve tag set that may be customized for the individual’s retrieval needs: {Milan, Italy, Gilberto, Peachy, Dad, Mom, Mipel, Trade, Show, Trip, Vacation, Business}. We’ve got the idea: Mom or Dad—they were at a trade show in Milan, Italy, on business. Gilberto and Peachy—they tagged along on the trip as a vacation. Here’s a different tag set that was applied across a whole vacation’s worth of photos: {Overseas trip 2006, Cinque Terre, Italy, Bologna, Milan, Rome, Greece, Santorini, England, Devon, Portsmouth, Isle of Wight}. It is hard to imagine the utility of these tags, even if these are indeed all the locations on the itinerary; because they have been applied across the entire photo set, the extra locations may be misleading descriptors for any individual photo.

What I mean to say is: tags can be a rich source of noise.

To really gain purchase on this problem, what we need is a head-to-head comparison of these three sources of evidence for retrieval, description, or any of the other purposes of extrinsic metadata.

So I counted. And counted. I looked for popular words, the ones lots of people used to describe the mosaic or the person spinning on the mosaic or the heel of the spinner. Table 4 lists the 20 most popular words used in tags, titles, and narratives. Words in red span all three lists; words in blue are in two out of three; words in black are only in one. The percentages are relative to the totals (tags, title words, or narrative words).

<table>
<thead>
<tr>
<th>20 most popular tags</th>
<th>20 most popular title words</th>
<th>20 most popular narrative words</th>
</tr>
</thead>
<tbody>
<tr>
<td>Milan</td>
<td>bull</td>
<td>bull</td>
</tr>
<tr>
<td>Italy</td>
<td>luck/lucky</td>
<td>luck/lucky</td>
</tr>
<tr>
<td>Milano</td>
<td>balls</td>
<td>spin/spinning</td>
</tr>
<tr>
<td>galleria/galeria</td>
<td>toro/torello/touro</td>
<td>good</td>
</tr>
<tr>
<td>bull</td>
<td>Milano</td>
<td>heel/heel</td>
</tr>
<tr>
<td>variants of Galleria Vittorio Emanuele II</td>
<td>spin/spinning</td>
<td>balls</td>
</tr>
<tr>
<td>Italia/Italian</td>
<td>Galleria Vittorio Emanuele (II)</td>
<td>Galleria Vittorio Emanuele (II)</td>
</tr>
<tr>
<td>toro/torello/touro</td>
<td>good</td>
<td>toro/torello/touro</td>
</tr>
<tr>
<td>europe/european</td>
<td>Milano</td>
<td>Milano</td>
</tr>
<tr>
<td>duomo</td>
<td>galleria/galeria</td>
<td>testicles</td>
</tr>
<tr>
<td>balls</td>
<td>step/stepping</td>
<td>three/3/three times/3x</td>
</tr>
<tr>
<td>travel/travelogy</td>
<td>palle</td>
<td>mosaic/mosaico</td>
</tr>
<tr>
<td>luck/lucky</td>
<td>testicles</td>
<td>turn/turning</td>
</tr>
<tr>
<td>Galleria Vittorio Emanuele (II)</td>
<td>Italy</td>
<td>step/stepping</td>
</tr>
<tr>
<td>mosaic/mosaico</td>
<td>mosaic/mosaico</td>
<td>Milano</td>
</tr>
<tr>
<td>palle</td>
<td>heel/heel</td>
<td>galleria/galeria</td>
</tr>
<tr>
<td>trip</td>
<td>fortuna</td>
<td>palle</td>
</tr>
<tr>
<td>vacation</td>
<td>three/3/three times/3x</td>
<td>tradition</td>
</tr>
<tr>
<td>spin/spinning</td>
<td>turn/turning</td>
<td>foot/feet</td>
</tr>
<tr>
<td>Lombardia</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| 60.1%                | 69.9                        | 85.5                           |
Table 4. Popular terms in tags, titles, and narratives

What we see is revealing. First, look at the totals. Only 60% of the tag words are accounted for in the list; 70% of the title words are accounted for; and more than 85% of the narrative words (not including stop words like articles) are accounted for.

What does this mean? In particular, what’s up with the other 40% of the tags? They’re tags that are low-frequency terms. Some are low frequency for the right reason: they’re relevant to the individual doing the tagging and not to the photo’s more public life in the Flickr database. For example, 5.7% of the tags are peoples’ names; these names are acknowledged as useful photo metadata. Others are less auspicious—non-standard variations of place names, unusual words that are tied in with the bull mosaic’s legend (feet, for example), and miscellaneous other words like summer and ribbon that seem to have simply popped into the tagger’s head when he or she drew a blank at tagging time. One tag was tellingly freeassociation.

It’s unlikely that these odd words will figure into future retrieval, either by the individual who assigned them or by someone like me.  

You’d think that this lack of uniformity would be true of the titles and narrative descriptions as much as it is for the tags. But—contrary to intuition—there’s more overlap in the terms used in the descriptions and titles than there is in the tags. Even words that are slightly more common in the tags—Italia, europe, duomo, travel, trip, vacation, and Lombardia—do not occur in the other lists. In fact, it is interesting to note that all of the title words occur in either the narrative, the tags, or both.

Tags are indeed miscellaneous, and that miscellany may make them less than useful.

It’s also telling that general place names are over-represented in the tags: Italy accounts for 11% of the total tags, and only 1.5% of title words. It figures into the narratives even less often than that. Certainly the term ‘Italy’ is too broad of a net to cast against the entire Flickr photo base to come up with our bull photo; it may even be too broad to use to retrieve a meaningful number of personal photos. Milan, used alone, is also doubtlessly too broad, although when used with other terms (such as bull or mosaic), it’s reasonably effective. Although Milan is the most popular tag, we also have to

---

8 Past experience tells me that over time, people do forget the existence of a significant portion of their photos; natural human forgetfulness—coupled with the unlikely tags—makes me think that these tags are unlikely to be useful for personal information retrieval. I have performed a number of field studies in which I accompany informants into basements, attics, guest rooms, closets, and storage areas—into the refuges of the dusty and forgotten—and mostly people are surprised not only by what they find, but also by what they don’t find. Our memories play cruel tricks on us.

9 Searching tags only, at the time of the original draft of this article, the word Italy retrieves almost 2.5 million photos, far too many to comb for any individual topic.

10 150,000 photos are tagged Milan (again, at the time of the original draft).
consider that Milan is often not meaningfully modified through other tags. For example, ‘Milan, Italy,’ is not appreciably more information than ‘Milan’.

The other place-related term that appears in all three lists is more specific than Milan. Although ‘Galleria Vittorio Emanuele’ is not as precise a description as one would like—it turns up 2,420 public photos in Flickr—it isn’t hopeless. 71 out of the 2,420 are of our bull. Unfortunately, to find them, one must page through 101 pages of thumbnails (24 thumbnails at a go). Even more unfortunately, there are only 3 in the first dozen pages and none in the first two. Most searchers will have given up before they reach the first positive result. Finally, for some reason, 26 of the photos of the bull mosaic are in the final 20 pages. I don’t know about you, but I know very few people sufficiently obsessive-compulsive to go through that many search results, especially given the low density of desired results.

In other words, I’m tired and dizzy from scanning so many images and doubt anyone else would do likewise.

By comparison, a query like ‘Milan bull’ turns up only one tenth of the photos of the query whose results we just exhausted. One-tenth! Of those 240 photos, 117 are what we’re after. Almost half. More than a third of the 322 photos possible. And 19 of the first 24 (that is, the first page of thumbnail results) are the desired photos.

Furthermore assigning general place names makes poor use of human labor in an era when GPS devices are cheap and relatively common and gazetteers can be used to fan out from the terms referring to specific places (Galleria Vittorio Emanuele) to more general areas (Milan, Italy, Europe).

Let’s isolate tags and terms related to derivable place (e.g. Italy) or time (e.g. summer, or more commonly a specific date) and see what happens. I’m distinguishing the most specific place names (all of the variants of galleria) from the more general ones, because they may be harder to assign on the basis of GPS data alone.

<table>
<thead>
<tr>
<th>% of general place names</th>
<th>% of specific place names</th>
<th>% of time</th>
<th>% of names of people</th>
</tr>
</thead>
<tbody>
<tr>
<td>tags</td>
<td>45.5 %</td>
<td>9.5 %</td>
<td>3.1 %</td>
</tr>
<tr>
<td>title terms</td>
<td>11.4 %</td>
<td>7.3 %</td>
<td>0.0 %</td>
</tr>
<tr>
<td>narrative terms</td>
<td>7.7 %</td>
<td>6.7 %</td>
<td>0.6 %</td>
</tr>
</tbody>
</table>

Table 5. Terms related to time and place

The message here is almost painful: a great proportion of user tags add little or no further information; as such, they don’t appear as often in narratives or titles. Personal names, which may be quite useful for finding photos among one’s own collection (especially over the long haul) are less well represented in all types of metadata, but are relatively similar in quantity.
Now here’s a property of tags that I find almost comical: they are seldom verbs, even if a verb is just the thing to characterize a photo. What’s unique about what tourists do when they visit the Galleria’s bull mosaic? They spin. In fact, if you type in Milan spin as your Flickr search terms, you pull up 94 results, 70 of which are pictures of our bull mosaic. 20 out of 24 results on the first page are on target.

Although spin and spinning make the top 20 list of tags, they are by no means commonly used terms; they are used less than 1% of the time (0.7%). That’s just 7 tags. On the other hand, spin makes up 4.8% and 9.5% of title and narrative terms. People just don’t seem to be thinking of tags as verbs.11

Have I convinced you that tags aren’t all they’ve cracked up to be? I hope I have, but nonetheless there’s a lingering fascination. Surely there’s something to be done about tags: we don’t want to just turn up our noses at Mr. Weinberger’s argument. They could be a compact and efficient way of describing pictures. After all, picture archiving is difficult. Witness Art Spiegelman’s fine graphical account in the New Yorker more than a dozen years ago; he described the difficult work of senior librarian Arthur Williams who curated the New York Public Library’s extensive picture collection for over 30 years12. Just how do you turn a library patron’s question, “I want a picture that conveys rough times ahead” into a photo of a three-masted schooner sailing into a storm?

In other words, knowing what we know about bull mosaics in Milan, what would make tags a more effective way of describing them?

First, it seems that it would be nice to encourage tag specificity. How come the same person who tags his photo {Milan, Italy} is able to title it Galleria Vittorio Emanuele II? Can we elicit the most specific place names possible using latitude/longitude pairs and fill in the rest using a gazetteer? Recognition is good enough to identify plausible human shapes in a photo. Can we elicit names? Those will come in handy too on down the line.

There’s plenty of evidence that many of our taggers know the story associated with the bull. How can we elicit the story and turn it into tags? Can we use an authority list of some sort to guide and normalize a user’s input? It is odd and telling that the terminology used in freeform input is more regular than the terms used in the tags. Narratives and even titles use verbs. How can we introduce verbs into the tag sets?

I find it odd that there’s been such a widespread and good-natured acceptance of the efficacy of tags. Yet—and maybe I’m being a Pollyanna here—there must be a way to turn them into the linguistic powerhouses they are advertised to be.

---

11 This is the time to think back to that exemplary tag set I asked you to remember earlier: {Italy, Italia, milan, milano, bull, good, luck, galleria, vittorio, emanuele, II, mall}. No verbs.
Where I live, there’s a local dairy, Berkeley Farms, that uses my favorite ad campaign ever. On the side of each Berkeley Farms milk carton, there’s a cow—a happy cow—and the lovely rhetorical question: Cows in Berkeley? Moo.

So now we can only say:

Bulls in Milan? Moo.

Turn around 3 times on your heel in Milan (*heel turn Milan* returns 8 photos, all of them of the bull mosaic) and you’ll find them.