



## UvA-DARE (Digital Academic Repository)

### Another take on tags? What tags tell

Simons, J.

**Publication date**  
2008

**Published in**  
Video Vortex reader: responses to YouTube

[Link to publication](#)

#### **Citation for published version (APA):**

Simons, J. (2008). Another take on tags? What tags tell. In G. Lovink, & S. Niederer (Eds.), *Video Vortex reader: responses to YouTube* (pp. 239-254). Institute of Network Cultures. [http://networkcultures.org/wpmu/portal/files/2008/10/vv\\_reader\\_small.pdf](http://networkcultures.org/wpmu/portal/files/2008/10/vv_reader_small.pdf)

#### **General rights**

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

#### **Disclaimer/Complaints regulations**

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

# ANOTHER TAKE ON TAGS? WHAT TAGS TELL

JAN SIMONS

## Problems with tags<sup>1</sup>

Tags and tag clouds are icons of what has become known as 'Web 2.0', the upgraded version of the Internet which promises the user: 'You control your own data'.<sup>2</sup> Tagging systems are considered as key instruments for this user control because they allow users to assign keywords of their own choosing to Internet resources of their own making as well as – at least in some cases – to objects produced by others. These tags are often primarily added for personal use, but in most so-called social network sites, tags are also accessible for other users. Because tagging systems allow users to freely choose and assign their own keywords, they are expected to liberate the users from the oppressive authority of pre-fixed, pre-defined, and often not readily intelligible vocabularies of prescribed top-down expert classification systems and taxonomies. In line with the guiding idea of Web 2.0 as the information infrastructure that facilitates the emergence of 'the wisdom of the crowds' – the Web 2.0 variety of what used to be called 'collective intelligence'<sup>3</sup> and nowadays best exemplified by the collectively produced web encyclopedia *Wikipedia* – the idea is that through the myriad interactions of individual users tagging systems will eventually make the tags chosen and assigned by individual users converge into a emergent bottom-up common vocabulary or *folksonomy*.<sup>4</sup>

But what would such a folksonomy look like? After all, users may apply keywords in different ways and tag resources for different purposes; tags are applied to a wide – if not wild - variety

- 
1. This article is based on a presentation at the conference *Videoortex: Responses to YouTube*, 18-19<sup>th</sup> of January in Amsterdam, Netherlands. Another version of this presentation was published as Jan Simons, 'Tag-elese or the Language of Tags', *Fibreiculture Journal*, Issue 12 'Models, Metamodels and Contemporary Media', 2008, [http://www.fibreiculture.org/issue12/issue\\_12\\_simons.html](http://www.fibreiculture.org/issue12/issue_12_simons.html)
  2. Tim O'Reilly. 'What is Web 2.0: Design Patterns and Business Models For The Next Generation of Software' (2005), <http://www.oreillynet.com/lpt/a/6228>
  3. Pierre Lévy. *Collective Intelligence: Mankind's Emerging World in Cyberspace*. Cambridge, Ma.: Perseus Books, 1997.
  4. Peter Merholz. 'Metadata for the Masses' (2004), <http://www.adaptivepath.com/ideas/essays/archives/000361print.php>; Peter Mika. 'Ontologies Are Us: A Unified Model of Social Networks and Semantics,' *Journal of Web Semantics: Sciences, Services and Agents on the World Wide Web* 5.1 (2007): 5-15; Clay Shirky. 'Ontology is Overrated: Categories, Links, and Tags', *Clay Shirky's Writings About the Internet: Economics & Culture, Media & Community*, Open Source (2005), [http://shirky.com/writings/ontology\\_overated.html](http://shirky.com/writings/ontology_overated.html); Thomas Vander Wal. 'Explaining and Showing Broad and Narrow Folksonomies', (2005), [http://www.personalinfocloud.com/2005/02/expaining\\_and\\_.html](http://www.personalinfocloud.com/2005/02/expaining_and_.html)

of objects (texts, blogs, bookmarks, photos, videos, music files, book titles, films, games, and what have you), and tagging systems themselves vary in design, ranging from 'blind' systems that let their users freely choose their tags to systems that show tags that other users assigned to the same or similar objects and gently recommend the user to choose one of these, and from 'narrow' systems that allow users to only tag their own documents to 'broad' systems that allow users to freely tag other users' objects as well.<sup>5</sup> These are so many incentives for very divergent, idiosyncratic, and inconsistent uses of tags and it is hard to see how anything like a controlled vocabulary could ever emerge from this highly uncontrolled labeling of what are often themselves already highly chaotic collections of heterogeneous objects.

It is hardly surprising, then, that the most discussed problems with free tagging systems are polysemy (does *music* refer to an audio file with music, a picture of a musical instrument, a file with a musical score?), homonymy (is *rock* used to tag a picture of a rock formation or a video or a rock concert?), synonymy (a Macintosh computer can be tagged with either *computer* or *apple*, and the latter can also be used to tag a piece of fruit, the former record company of *The Beatles*, or in combination with *big* to refer to New York City), and differences in levels of categorisation (a pet can be tagged with *cat*, *feline* or *animal*). To these problems one could add spelling and orthography. Many users of *Flickr*, for instance, tag pictures of New York as *new york* or *new york city* without realising that Flickr's tagging system interprets isolated strings of symbols as separate tags. The result is on the one hand the unexpected appearance of the city of York in Flickr's 'all time most popular tags', together with the lexical item *San* of which complements like *Francisco*, *José*, or *Bernardino* didn't make it to this distinguished list, while on the other hand New York (and other cities with composite names) is tagged in at least three different ways (*newyork*, *newyorkcity*, *nyc*). Other problems are unlikely compounds (*TimBernersLee*, *sometaihurts*, *handsclawsandallkindsofpaws*), personal tags (*mydog*, *me*, *natasja*) or one-offs (*billybobsdog*).<sup>6</sup>

At closer inspection, however, it turns out that polysemy, synonymy, homonymy, and levels of categorisation only scratch the surface of the semantic problems with tags. Is a picture tagged with *england* a picture *from* or *about* England, or is it simply a picture taken somewhere *in* England? Do tags like *red*, *green*, *blue* refer to properties of the objects in a picture, do they refer to salient properties of a photograph itself (e.g. the use of filters, or the application of post-production techniques), or even to a property that makes a picture eligible for admittance to a group or 'pool' of pictures? These problems do not arise from any ambiguity of the meaning of words like *England*, *red*, *green* or *blue* – there is nothing ambiguous about

5. Vander Wal, 'Explaining and Showing Broad and Narrow Folksonomies'; Cameron Marlow, et al. 'HT06, Tagging Paper, Taxonomy, Flickr, Academic Article, ToRead', *Proceedings of the Seventeenth Conference on Hypertext and Hypermedia*. New York: Association for Computing Machinery (ACM), 2006. Also available from <http://www.danah.org/papers/Hypertext2006.pdf>; Scott Golder and Bernardo A. Huberman. 'The Structure of Collaborative Tagging Systems', *Journal of Information Science* 32.2 (2006): 198-208.
6. Adam Mathes, 'Folksonomies – Cooperative Classification and Communication Through Metadata' (2004), [http://blog.namics.com/2005/Folksonomies\\_Cooperative\\_Classification.pdf](http://blog.namics.com/2005/Folksonomies_Cooperative_Classification.pdf); Marieke Guy and Emma Tonkin, 'Folksonomies: Tidying Up Tags?', *D-Lib Magazine* 12.1 (2006), <http://www.dlib.org/dlib/january06/guy/01guy.html>

the meanings of these words in themselves – but rather from the different relations these tags entertain with the objects they are used for.

Of course, one could argue that these ambiguities dissolve the very moment these tags are seen in connection with the objects they label. But the point of tagging is precisely to make objects retrievable either for later use by the producer/owner of the object or for other users if the objects are stored on social network sites like *Flickr*, *YouTube*, *Last.fm*, *Del.icio.us* and others. As *Flickr* explains, tags are 'like a keyword or category label. Tags help you find photos and videos which have something in common'.<sup>7</sup> But tags can only accomplish this if their meanings are largely independent of the particular objects they label, because if one first needs to match an object with a tag in order to disambiguate its meaning, tags would be rather poor instruments for finding objects in the first place. Using a tag for a search would be like firing a shot in the dark. And such indeed may be the sad truth about many tagging systems....

### Problems with tags

But the problems with tags may result from the particular ways in which they are thought and theorised and the particular expectations that are brought to bear on them from these theoretical perspectives rather than from some inherent properties of tags themselves. As the end product of the myriad interactions of individual tagging practices a folksonomy is eventually nothing but a democratically – or rather, collectively – defined taxonomy. Just like classical taxonomies, folksonomies use tags as labels to identify objects for purposes of classification, indexation, and retrieval. Whether the meanings of these labels are prescribed and imposed by expert authorities or result from a consensus that has emerged spontaneously from the free practices of the multitudes of users, both expert terms and democratic tags are expected to reliably and unambiguously identify and retrieve the members of the categories they are supposed to cover. This unavoidably means that once the tagging practices of the multitudes have settled into a more or less stable and controlled vocabulary, newcomers to a folksonomy will have to conform to the established consensus and 'learn' the meanings and proper uses a tagging community has already assigned to their tags which is in principle not that much different from learning the terms of an expert taxonomy (it is very well conceivable that a folksonomy will be published as a 'wiki-dictionary'). Proposals for the remediation of the flaws of tagging systems already indicate solutions pointing in that direction.<sup>8</sup>

The problem with these approaches to folksonomies is not so much the unrealistic expectation that some stable vocabulary will eventually emerge from tagging, but rather that their view on tagging itself is tainted by the very taxonomies they want to provide an alternative for. This view could itself be tagged as what Christian Metz, one of the founders of French film semiotics, once called a 'FIDO-fido'-view of language because it treats lexical items in general and nouns in particular as labels that 'name' objects like the proper noun *Fido* 'rigor-

7. Flickr, <http://flickr.com/photos/tag>

8. Guy and Tonkin, 'Folksonomies: Tidying Up Tags?'; Mathes, 'Folksonomies

ously' identifies the dog that goes by that name.<sup>9</sup> This view is, of course, encouraged by the very purposes taxonomies and folksonomies are designed to serve, which is the classification and identification of particular objects. The 'FIDO-fido' view is further encouraged by the very design of taxonomies and tagging systems: most classificatory systems, including tagging systems, allow users to use only separate and discrete lexical items as labels. Tagging systems, that is, force users to treat words as single and isolated items with which they must capture and 'name' a salient or significant aspect of the target object. This accounts for the heavy predominance of nouns in tagging systems, since it is part of the semantics of nouns that they typically denote discrete 'bounded entities' and because unlike verbs, adjectives, adverbs or prepositions nouns typically do not require other lexical items as syntactic complements.<sup>10</sup> Moreover, tagging systems are typically a-synchronous and asymmetrical systems for 'communication through metadata': taggers do not interact with each other in 'real time' and on a person-to-person basis but they get feedback from tagging systems that presents them with the aggregate results of previous tagging practices. The system, however, only samples tags on the basis of the frequency of their use but it is blind for the particular context in which tags were used and has no clue as to the particular meanings tags might have had in those contexts, let alone being able to provide prospective taggers with such clues. Lexical items, that is, appear in tagging systems deprived from any semantic, syntactic or pragmatic context which might help to 'ground' their meaning.

These factors promote a 'dictionary' approach of lexical items and a 'picture' view of language, in which words have well circumscribed and context independent meanings and in which linguistic expressions are supposed to 'mirror' real world objects or states of affairs. In other words, tagging systems favor the classical positivist view of language in which words should ideally have definite meanings and according to which the meaning of an expression consists of its 'truth conditions'.<sup>11</sup> In this respect proponents of folksonomies only disagree with (what they perceive as) authoritarian taxonomists about how and when tags will acquire their 'definite meanings': top-down or bottom-up, before or after tagging.<sup>12</sup> However, the concern for a reliable and stable classification system might obscure what is actually going on in tagging practices and prevent seeing with clarity the emergence of an order that

- 
9. Christian Metz, 'Le Cinéma: Langue ou Langage?', in *Essais sur la Signification du Cinéma*. Tome 1. Paris: Klincksieck, 1983, p. 68. Metz borrowed the label 'FIDO-fido theory' from the American philosopher G. Ryle.
  10. See Ronald W. Langacker, *Cognitive Grammar: A Basic Introduction*. Oxford: Oxford University Press, 2008; Adele E. Goldberg, *Constructions at Work: The Nature of Generalisation in Language*. Oxford: Oxford University Press, 2006.
  11. See A.J. Ayer, *Language, Truth and Logic*. Harmondsworth, UK: Penguin Books, 1987; Ludwig Wittgenstein. *Tractatus Logico Philosophicus*. London and New York: Routledge, 2001 [1922].
  12. This is, for instance, what Clay Shirky has to say about the 'filtering' process in folksonomies: 'Similarly, the idea that the categorization is done after things are tagged is incredibly foreign to cataloguers. Much of the expense of existing catalogue systems is in trying to prevent one-off categories. With tagging, what you say is 'As long as a lot of people are tagging any given link, the rare tags can be used or ignored, as the user likes. We won't even have to expend the cost to prevent people from using them. We'll just help other users ignore them if they want to'. Clay Shirky, 'Ontology is overrated'.

the 'wisdom of the crowds' actually has introduced in the seemingly 'feral', 'out-of-control' hypertext structure of tagging systems.<sup>13</sup> Folksonomists might have been looking for order at the wrong place.

### What is the problem with tags?

If one takes the brief of Web 2.0, 'You control your own data', seriously it does not make much sense to approach what users actually *do* with their data from the point of view of some desired end result (e.g., a 'controlled vocabulary') only to come up with remedies for perceived problems (e.g., polysemy, homonymy, synonymy, etc.) that get in the way of achieving the desired outcome.<sup>14</sup> Rather, one should try and approach users' practices from a user's perspective to try and find out what they actually do and how they proceed when they attempt to get a grip on their data by tagging them. Although this seems to be an impossible task, given the massive numbers and the geographical dispersion of taggers, taking a user's perspective does not necessarily require that the researchers sits down next to a tagger in order to 'examine user's motivations when adding tags, see why they decide on particular words, observe how many tags they add and compare how the same items are classified by different users'.<sup>15</sup> The higher-level order that emerges from the myriad interactions of the components of complex systems typically displays properties and behaviours that are not reducible to or predictable from those interactions (nor are those properties and behaviours the results of bottom-up, 'democratically' made choices). Higher-level orders tend to 'transcend' the behaviours, perceptions and awareness of the individuals who participate in the lower-level activities that give rise to higher-level orders, which makes it very unlikely that research methods that focus on the decisions of individual users are capable of grasping even the onset of higher-level complexity.

Languages, for instance, are good examples of complex systems that emerged from countless usage events through which units, schemas and constructions of different levels of gen-

- 
13. Jill Walker, 'Feral Hypertext: When Hypertext Literature Escapes Control', *Proceedings of the Sixteenth ACM Conference on Hypertext and Hypermedia, Hypertext '05*. New York: ACM Press, 2005. See also Marlow et al. 'HT06, Tagging Paper, Taxonomy, Flickr, Academic Article, ToRead'.
  14. The term Web 2.0 has itself been tagged in various and not always synonymous ways. The *Webopedia*, for instance, offers the following definition of Web 2.0: 'Web 2.0 is the term given to describe a second generation of the *World Wide Web* that is focused on the ability for people to collaborate and share information online. Web 2.0 basically refers to the transition from static HTML Web pages to a more dynamic Web that is more organised and is based on serving Web applications to users. Other improved functionality of Web 2.0 includes open communication with an emphasis on Web-based communities of users, and more open sharing of information. Over time Web 2.0 has been used more as a marketing term than a computer-science-based term. Blogs, wikis, and Web services are all seen as components of Web 2.0.' This definition hovers between a technical definition and a definition in terms of the social uses and practices afforded by applications served to users. The observation that 'Web 2.0' has over time become a marketing term is nothing less than a recommendation not to dump the term altogether. *Webopedia*, [http://www.webopedia.com/TERM/W/Web\\_2\\_point\\_0.html](http://www.webopedia.com/TERM/W/Web_2_point_0.html)
  15. Marieke Guy and Emma Tonkin, 'Tidying up Tags?' For practical reasons, Guy and Tonkin do not apply an ethnographic methodology themselves, since 'such studies take time and resources'.

erality became entrenched and conventionalised.<sup>16</sup> Native speakers must have a ‘knowledge’ of their language, but since for most native speakers the grammar of their native language is not consciously available, and no single native speaker has a full grasp of his or her native language – even professional linguists still have not managed to deliver a full account of English, the world’s most studied language system. In order to study the grammar of a language, linguists therefore do not examine native speaker’s motivations, choices and decisions when they speak. They study the language system through the traces it leaves in the actual and possible utterances native speakers produce and understand. In order to discover the system that might emerge from countless individual tagging events, one had better study the traces of these events rather than the particular motivations and choices that led to the event.

If tags and tagging practices can be considered as ‘a largely user-driven adaptation of natural language for indexing purposes’, the characteristics of such a ‘sublanguage’ can be studied ‘as a corpus of interest to linguistics’ and not only ‘as a keyword corpus in need of filtering’, as has been done in most, if not all, studies of tagging systems.<sup>17</sup> Assuming that for their tagging practices taggers tap on their common, everyday knowledge of language rather than being driven by the concerns of taxonomists, archivists and other professional indexers – and isn’t that the point of Web 2.0? – tags can be considered as utterances or speech acts. But since tags are devoid of syntactic, semantic and pragmatic context because of the limitations tagging systems impose on users, they seem to embody the paradoxical phrase Christian Metz once coined to characterise film: a language without grammar (*langue sans langue*).<sup>18</sup> How to get out of this conundrum?

Tagging systems may deprive users of the means to express syntactic and semantic relationships, but that does not mean that tagging systems make the users’ semantic and syntactic knowledge inoperative. On the contrary, there is evidence that this knowledge directs tagging practices. A well-known strategy to circumvent the obligation to use single words as tags is the creation of compound words consisting of more terms or mixtures of languages (e.g., *hardrockcalling*, *macysfireworks*, *happybirthdayamerica*, *pyramidstage*, *bisousbacikussekisses*).<sup>19</sup> Moreover, most users of Flickr and other sites with tagging systems follow the recommendation to use two or more tags in order to increase the retrievability of their data. Of course, users may have many incentives to tag their data with often many different and unrelated tags but it is not to be excluded that one of the incentives is the desire to contextualise tags in one way or another. But then again, precisely because users may have many motives for choosing tags – to attract as many other users as possible, to deceive a site’s watchdogs who guard against indecent, offensive, politically correct or copyrighted content, to advertise their services, or for sheer personal use (*mydog*), and so on – a ‘gram-

16. Ronald W. Langacker, *Cognitive Grammar*, pp. 216 ff; Adele E. Goldberg, *Constructions at Work*, pp. 63 ff.

17. Emma Tonkin, ‘Between Symbol and Language-In-Use’, Bich-Liên Doan, Jose Joemon and Massimo Melucci (eds.) *Proceedings of the 2<sup>nd</sup> International Workshop on Context-Based Information Retrieval*. Roskilde, Dk: Roskilde University, 2007, p. 116-117.

18. Christian Metz, ‘Cinéma: Langue ou Langage?’, p. 70.

19. Retrieved from *Flickr’s* ‘hot tags’ on July 6, 2008.

mar’ of ‘tag-elese’ – if there is such a thing – should not be ramified at the level of individual usage events but rather at a higher, aggregate level of these events.

The strongest indication that ‘ordinary’ taggers tap into the resources of everyday language are perhaps the very phenomena, that according to their critics, indicate that folksonomies require filtering devices: polysemy, homonymy, synonymy, different levels of categorisation, and the use of ‘non-dictionary’ expressions such as compound words. In natural languages, polysemy, homonymy, synonymy and ambiguities of all sorts are not deviant, but default. Unlike the terms in taxonomies, dictionaries, or encyclopedias, that are typically defined and sanctioned by expert (individual as well as collective) authorities, words in natural languages do not ‘have’ meanings but function as *cues* to meaning urging the language user to search an intricate and open-ended network of senses that may vary from novel interpretations through incipient sense to established conventional or ‘central’ meanings.<sup>20</sup> Syntactic, semantic and pragmatic context guide language users on their search for a suitable but always provisional and hypothetical interpretation for the cues in question, but when such contextual cues are prohibited, as in tagging systems, cues become incentives for a seemingly uncontrollable proliferation of meanings. However, in language usage lexical items themselves are not only cues to the semantic content but they usually also convey cues to syntactic information, semantic relations and possible pragmatic usage content. Could it be possible that this information is somehow preserved and recoverable at a higher aggregate level of tag usage events?

### A Walk In The Clouds

Fortunately, the web itself provides the aggregate data that are relevant for a search for system in the chaos of tags. The photo – and now also video – hosting website Flickr, for instance, provides a so-called tag cloud in which the 150 all time most popular (i.e. most frequently used) tags are listed alphabetically with their relative frequency of use being represented by font size (the bigger the font, the more popular the tag). Flickr offers its users a ‘blind’ and ‘narrow’ tagging system, which means that users are completely free to choose tags (Flickr does not come up with recommendations), and users are allowed to only tag their own images or videos unless they have gotten explicit permission to tag photos or videos of other users from the owners of those photos or videos. It does not seem unreasonable to assume that Flickr’s tag cloud is a quite reliable representation of choices made by taggers when left to their own devices.

20. Ronald W. Langacker, *Cognitive Grammar*, p. 38-39; Jeffrey L. Elman, ‘An Alternative View of the Mental Lexicon’, *Trends in Cognitive Sciences*, 8.7 (July 2004): 306; Gunther Kress, *Literacy in the New Media Age*. London and New York: Routledge, 2003, p. 39; Vyvyan Evans, ‘Lexical Concepts, Cognitive Models and Meaning-Construction’, *Cognitive Linguistics* 17.4 (2006): 493.



Figure 1: Flickr tag cloud 2007.

The tag cloud, of course, does not provide an insight in the particular tagging practices of individual users. First of all, these 150 tags are only a tiny part of the huge amount of actually used tags. The tag cloud represents the relatively few tags that are used by (very) many users, but it leaves the many tags used by a smaller amount of users and the huge number of tags used by only one or two users out of the picture. The distribution of tags, that is, follows a power law according to which ‘the rich get richer’: the tag cloud itself, for one thing, provides the most popular tags with a higher visibility than then the massive amount of tags that didn’t make it into this cloud and might thus function as a kind of a recommendation system (a tagger who wants to increase the ‘visibility’ of his photos had better choose one of the all time popular tags than come up with a personal and idiosyncratic tag, for instance).<sup>21</sup> Under the reign of a power law, devices like a tag cloud are bound to instigate positive feedback loops and the relative stability of Flickr’s tag cloud over a longer period of time seems to confirm this.

This, however, does not say very much about the meaningfulness of tags, since meaningfulness is a matter for the individual tagger or groups of taggers to decide: there is no reason to assume, for instance, that under the pressure of a power law a folksonomy will eventually converge on a relatively limited set of tags with consensually agreed upon meanings, since the power of a power law in environments like the web is precisely that it allows room for particular individual points of view or niche interests in the so-called ‘long tail’.<sup>22</sup>

21. Adam Mathes, ‘Folksonomies’, p. 11; Marieke Guy and Emma Tonkin. ‘Tidying up Tags?’.

22. Mathes, for instance, suggests that examining power law distributions of tags ‘could give a better indication of whether a folksonomy converges on terms and foster consensus, or if as the user base grows the vocabulary grows at a more even rate, and the distribution of terms flattens, perhaps indicating less agreement.’ Mathes. ‘Folksonomies’, p. 11. Shirky, on the other hand, argues that the ‘market logic’ that dominates tagging practices ‘allows many distinct points of view to co-exist, because it allows individuals to preserve their point of view, even in the face of general disagreement’. Shirky, ‘Ontologies are overrated’. See also Chris Anderson, *The Long Tail: How Endless Choice Is Creating Unlimited Demand*. London: Random House Business Books, 2006.

The tags in the tag cloud, moreover, are selected merely on the basis of their frequency of use. The algorithms that sample and rank the tags on Flickr are blind to the content of the pictures tagged by these tags, as well as to the other tags with which they might co-occur: in this respect, the tags in the tag clouds are nothing but a bunch of de-contextualised clues. The only relationship between the tags in the tag cloud is one of co-occurrence in the tag cloud itself, that is, but not necessarily in any of the data tagged by them.<sup>23</sup>

Although this may seem to make the tag cloud look like a random collection of tags that only share their high ranking in Flickr’s popularity pool, it also, paradoxically perhaps, makes the task of searching for some system in the cloud easier. First of all, since the 150 all time popular tags have been severed from the pictures for which they were used, the particular content of those pictures is of some marginal interest: the tags in the cloud do not refer to any picture or any item of their content in particular. The semantics of these tags can only be interpreted at the most abstract or general level (which is not something like a dictionary meaning but rather a very schematic or central or prototypical meaning). This also means that whatever patterns are found in Flickr’s tag cloud can also be found in other tagging systems, since the found patterns are not dependent on the particular nature, modality and content of the tagged items (although this is, at the current state of research, a hypothesis that needs to be empirically tested).

Second, the tag cloud might be representative in more than one sense. Apart from representing the ‘most popular’ tags – the metadata most users have ‘voted for’ – patterns in the tag cloud, the ‘short head’ of the power law that governs the quantitative distribution of Flickr’s tags, are more than likely similar to patterns further up in its ‘long tail’. There is, for instance, a power law distribution of tags *within* the ‘short head’ itself (see Figure 1): only a few tags have been used more than 3500000 times, a greater number of tags have been used between 1750000 and 3500000 times, whereas the vast majority of these 150 tags has been used less than 1750000 times. Since power laws are fractal – ‘which is to say that no matter how far you zoom in they still look like power laws’ which means ‘that the Long Tail is made of many mini-tails’<sup>24</sup> – it is not unreasonable to assume that relationships found between tags in the short head will be found among tags in the ‘mini-tails’ that build up the long tail (again a hypothesis in need of empirical investigation).

### Tagging the tags

These assumptions made – for present purposes, the content of the tagged items is of marginal interest and under a power law the patterns found in one segment are ‘self-similar’ to patterns in any other segment – what pattern or patterns are there in Flickr’s tag cloud? A ranking on the basis of frequency of use does not yield more information than the tag cloud

23. It is not even clear that the ‘co-occurrence hypothesis’, according to which similar words are preferentially used for similar items, applies to the tag cloud, since the tag cloud itself gives no ‘cue’ to which of what kind of items the tags are attached. See Emma Tonkin, ‘Between Symbol and Language-In-Use’, p. 116.

24. Chris Anderson, *The Long Tail*, p. 139.

itself already provides, except maybe the working of a power law.<sup>25</sup> A better way to proceed is to try and classify the tags themselves. As was to be expected, for reasons already explained above, an overwhelming majority of 126 out of 145 tags consists of nouns, the remaining tags being distributed over adjectives (12), one verb (*hiking* – and one might question whether this progressive should not be classified as a modifier (adjective) or maybe even as a nominalization (noun)), one personal pronoun (*me*), one cipher (*07*), and some hard to classify tags such as *bw*, *la*, *dc*, *de* and *san*. The latter four are most likely components of proper names like *Los Angeles*, *Washington DC*, *San Francisco* and *Rio de Janeiro*, whereas the former probably is an abbreviation of *black-and-white*. The tag *new*, here classified as an adjective, is almost certainly also used as part of the proper name *New York* which makes it a component of a noun rather than an adjective.

A closer look at the noun tags learns that 44 out of 126 nouns (35%) are proper nouns that refer to continents, countries, regions, states or cities (USA states and cities by far outnumbering the rest of the world). Besides these 47 proper nouns there are some 15 tags such as *beach*, *sea*, *camping*, *river*, *mountain*, *zoo* that refer to locations as well, which brings the number of locative nouns up to about 62, which is almost half of the nouns in the tag cloud.

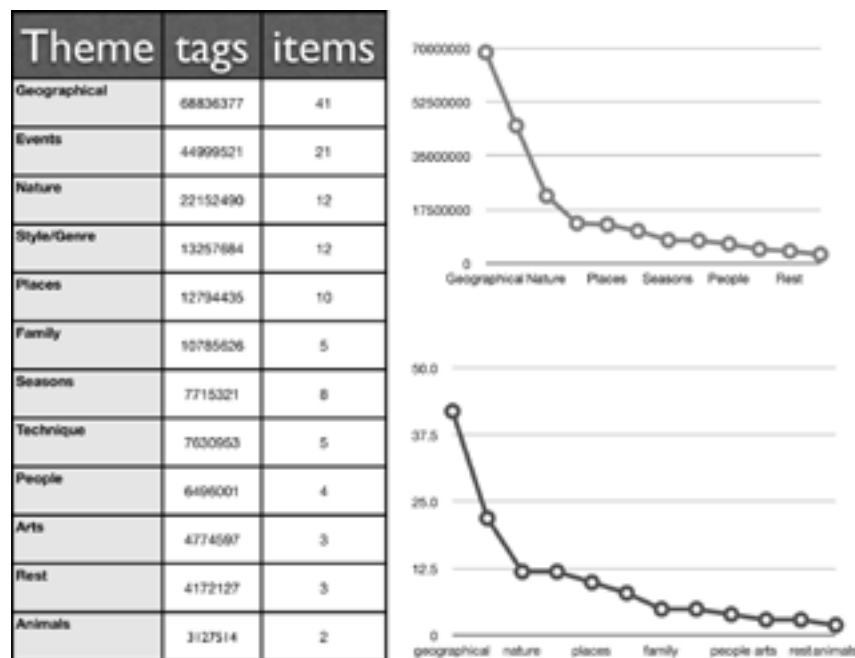


Figure 2: Table with tag categories.

25. In order to uncover the working of a power law one needs to have the exact figures of the frequency of use of tags, and plot these in a graph. Although the tag cloud itself does not provide these figures, they can be easily obtained by clicking on the tags. The page which opens displays the exact number of 'uploads' tagged with that term.

Apart from these locative tags there are also 15 out of 126 nouns (12 per cent) that refer to moments in time (e.g. *august*, *birthday*, *autumn*, *holiday*, *Christmas*). This relatively low amount of temporal tags can be partly explained by a technological feature of contemporary digital cameras that automatically register the time and date the picture was taken. These meta-data are automatically uploaded to Flickr together with other meta-data such as the type of camera the picture was taken with. Since these meta-data are displayed together with the tags added by users themselves every picture is automatically tagged with temporal information, which makes temporal tags provided by the users redundant. But more importantly, a look at a random sample of pictures tagged with temporal tags reveals that most of them show activities that are typical for the season, period or moments mentioned by the tags. Instead of referring to particular moments or periods in time, then, these tags metonymically refer to events that characterise the mentioned episodes. In that respect, temporal tags are often on a par with the 11 tags that explicitly refer to events or activities such as *festival*, *concert*, *trip*, or *travel*. However, the same goes for locative tags as well: these can also refer to a site as the subject matter of a photograph as well as to activities or events that took place at those sites. Like temporal tags, locative tags can also metonymically refer to activities and events.

The near absence of verbs seems to be largely compensated for either by nouns that refer directly to events or by temporal and locative nouns that metonymically refer to events. This is consistent with the fact that events are often metaphorically talked about as objects, as in *going to the concert*, *visiting a festival*, or *witnessing an explosion*.<sup>26</sup> Note that this polysemy of temporal and locative tags is a very common feature in natural language: *Christmas* can mean 'the most wonderful *time of the year*' but also the '*celebration of the birth of Christ*', the traditional '*family-get-together*' or whatever might be a typical activity for the speaker to undertake on the 25<sup>th</sup> of December.

Eleven of the 145 tags in the cloud are adjectives (7.5 per cent), seven of which are colour terms (that could also be classified as colour names and be counted as nouns, e.g. *blue*, *black*, *red*, *green*, *yellow*, *white*). Of the remaining adjectives one refers to a property of the picture (*geotagged*), one to a particular device with which the photo was taken (*macro*), another to a salient feature of a picture's content (*urban*), and, as already mentioned, one is rather to be considered as part of a proper name (*new* as component of *new york*). Except for the last one, all of the adjectives can be grouped as a cluster that pertains to the style or the genre to which the picture belongs.

Granted that at least a number of temporal and locative tags can be classified as events as well, some interesting patterns starts to emerge. First of all, as was to be expected in a set

26. See Georg Lakoff and Mark Johnson, *Metaphors We Live By*, Chicago: Chicago University Press, 1980, pp. 30-31; Jeffrey M. Zacks and Barbara Tversky, 'Event Structure in Perception and Conception', *Psychological Bulletin* 127.1 (2001): 7. De-verbalization does not magically turn a process into a thing but it signals that the process is being cognitively construed and linguistically profiled as a 'bounded entity'. It might well be that the static nature of photos together with the syntactic limitations imposed by the tagging system encourage a non-processual representation of events. See Ronald W. Langacker, *Cognitive Grammar*, p. 17-118.

governed by a power law, there is a 'short head' within the 'short head' of Flickr's tags as well, consisting of clusters of temporal, locative, and event tags that together make up about 69 per cent of the tag cloud. The remaining 46 tags (31 per cent) in this short head's 'long tail' can be distributed over 8 more categories: *nature* (12), *style/genre* (12: e.g., *portrait, landscape, urban, night, color*), *family and friends* (4: *baby, kids, family, friends*), *technique* (5: *cameraphone, canon, nikon, macro, film*), *people* (4: *people, me, portrait, girl*), *arts* (3: *architecture, art, graffiti*), *animals* (3: *cat, cats, dog*) and *rest* (3: *new, de, san*). Second, and more interesting for the purposes of this research, the tags within the 'short head within the short head' pertain to times, places, and events. Intuitively this seems to make sense: whatever a photograph depicts, it has to have taken place or have been located at some place at some time. Time, place and events are the basic components of what in human experience constitutes a scene and these basic experiential components are reflected in the basic argument structure of language, in which time, place, and event constitute a 'nuclear' argument structure.<sup>27</sup>

### Tag-elese

In natural language, the 'argument roles' that fill slots in a semantic argument structure like *agent, patient, instrument, theme*, etc., typically correspond to the participant roles that are lexically expressed in sentences as subject, object, indirect object or prepositional phrases. In the construction of argument structures and the organisation of syntactic structures the verb is pivotal, because it projects the roles that are involved in the process it denotes.<sup>28</sup> However, since the limitations of the tagging system prohibit the expression of a core argument structure with corresponding syntactic participant roles in subject-verb-object-indirect structure, the system leaves taggers with no other option than to subject all argument roles to a process of nominalisation and to express them as nouns. Core arguments such as time, place, and process are then lexically expressed as the temporal, locative and event tags that jointly make up the 'short head' within the tag cloud, whereas argument roles as *agent, patient, instrument*, etc. are 'demoted' to so-called 'satellite arguments' like 'additional participants' in the tag cloud's long tail.<sup>29</sup> Since nouns do not take subjects, the argument role of *agent* cannot take the corresponding syntactic participant role of subject, and no longer capable of taking on the agentive role in process, it must adopt the argument role of *theme*, which is then 'fitted' into the role of one of the 'satellite arguments' such as 'additional participants'. These satellite argument roles are typically filled with tags from the clusters *family and friends, people, and animals*.

27. Ronald Langacker, *Foundations of Cognitive Grammar*. Vol. 2: Descriptive Applications. Stanford, Cal.: Stanford University Press, 1991, pp. 294-295; Adele E. Goldberg, *Constructions: A Construction Grammar Approach to Argument Structure*. Chicago: The University of Chicago Press, 1995, p. 39.

28. Adele E. Goldberg, *Constructions at Work*, pp. 39-40.

29. See Anna Siewierska, *Functional Grammar*, London and New York: Routledge, 1991, p. 55, p. 72. In this respect, tagging systems force taggers to adopt expressive strategies that are remarkably similar to those used by the creators and speakers of pidgin languages (as was also observed in passing by Marieke Guy and Emma Tonkin, 'Tidying Up Tags?'). See Jan Simons, 'Tag-elese Or The Language of Tags'.

But what about the remaining clusters like *style/genre, technique, nature, and arts*? These seem to be tags that are situated at what might be called the level of a meta-discourse on the photographic act itself. After all, whatever event or state-of-affairs a picture depicts, a picture's content is always the result of the single event that is presupposed by every photograph which is the act of making the photograph itself. This meta-event requires itself an argument structure with corresponding participant roles: an *agent* (the photographer), a *patient* (the photographed persons, animals, or scenes) or *theme* (a state-of-affairs), an *instrument* (the equipment with which the photograph was taken), *manner* (the particular style or genre), and, of course, the *time* and *place* of the photographic event itself. Some of these meta-arguments are strictly speaking not part of the tags in the tag cloud, but they do appear in the meta-data that are automatically 'captured' and 'rendered' by Flickr: the (account) name of the owner/producer of the photo, the camera with which the photo was taken and the date and time at which the picture was taken. But as is to be expected on a site dedicated to photography, users may consider aspects of the making of the photograph as important as the content of the picture – or rather, they may consider the particular techniques, stylistic features, technical equipment or photographic skills with which the picture was taken as the picture's actual content. For these photographers, subjects like 'nature', 'landscapes', 'people' (*girls*, not *boys* who are remarkably absent from the tag cloud), or 'urban' scenes are occasions for exercising and displaying their photographic skills rather than being of interest in themselves. These meta-argument roles are lexically expressed in tag clusters such as *technique, style/genre, nature, and art*. But since photographic techniques and skills can be exercised on any subject, there is nothing that prevents these meta-tags being juxtaposed to tags pertaining to a photograph's content matter.

<b>ARGUMENT STRUCTURE:</b>					
TIME PLACE AGENT <- EVENT -> PATIENT/THEME					
<b>PARTICIPANT STRUCTURE:</b>					
	TIME PLACE [ TIME PLACE	SUBJECT NOUN [EVENT]	VERB NOUN [EVENT]	OBJECT/INDIRECT NOUN [OBLIQUE OBJECT] .....	OBJECT PP
<b>TAG-STRUCTURE:</b>					
META-TAGS:					
TIME/DATE + OWNER/PRODUCER + (EVENT) + CAMERA TAKING A PICTURE					
USER TAGS:					
TIME LOCATION + EVENT + PARTICIPANTS + INSTRUMENT + MANNER TEMPORAL EVENT FAMILY/FRIENDS TECHNIQUE STYLE/GENRE LOCATIVE PEOPLE ANIMALS					

Figure 3: Table with tag structure.

Flickr's tag cloud turns out to be governed by a remarkably stable structure, which is the argument structure familiar from the grammar of natural languages. Although this is not in itself surprising, since taggers, who are not professional or expert archivists, indexers or taxonomists have not much else to go by than their native knowledge of language, it is remarkable that this argument structure emerges in a system that deprives its users of the usual expressive means of a natural language. It is no less remarkable that this underlying argument



structure is not the result of a process of induction or abstraction from individual tagging practices of individual users. It is even highly unlikely that a full-fledged argument structure can be reconstructed from any particular set of tags attached to a unique photograph. Rather, the tag cloud displays properties that emerge from the aggregate interactions of literally millions of users, without any of them consciously applying this argument structure.

On the other hand, if it is true that power laws are fractal, this argument structure should at least partially be operative on every scale of the phenomenon it governs. This means that at least vestiges – or rudimentary parts – of the argument structure should be minimally recoverable in tagging practices at any point of the long tail. This certainly needs further investigation.

### Tag-elese or the Purloined Language

If there is an order in tags, it is certainly not an incipient order of a 'controlled vocabulary' the proponents of folksonomies expect – or hope – to emerge, but neither are folksonomies as 'feral' as critics of folksonomies fear. The patterns that seem to govern tagging practices are remarkably similar to those that govern everyday language usage: somehow, users appear to be trying to contextualise tags in order to 'pin' down their meaning, and they do this with the same cognitive and linguistic resources they have at their disposal in their everyday use of language. These semantic, syntactic and pragmatic resources, however, are not designed to make meanings of lexical items converge into some consensually achieved or 'democratically voted' controlled vocabulary, but rather to help speakers and listeners to choose a particular interpretation of a linguistic utterance in a particular usage event. One might argue that the grammar of a language is designed not to abolish but rather to preserve polysemy, synonymy, homonymy and other sources of ambiguity, because these are among the properties that provide languages with their enormous and vital flexibility.

Since taggers tap into the same cognitive and linguistic resources that allow for the impressive flexibility and adaptability of language, it is very unlikely that tagging practices will eventually converge in something like a controlled vocabulary. Tag-elese is not a 'language without a grammar', but its grammar is largely concealed – or 'repressed' as a Freudian would say – by the very design of tagging systems and – it should be admitted – by the very purposes proponents of folksonomies had in mind for tagging practices. Nevertheless, as the 'purloined letter' in Poe's famous story, the grammar of tag-elese has been staring us in the face all the time while we were looking for it at the wrong place.

### REFERENCES

- Anderson, Chris. *The Long Tail: How Endless Choice Is Creating Unlimited Demand*. London: Random House Business Books, 2006.
- Ayer, A.J. *Language, Truth and Logic*. Harmondsworth, UK: Penguin Books, 1987.
- Elman, Jeffrey L. 'An Alternative View of the Mental Lexicon', *Trends in Cognitive Sciences*. Vol. 8, no. 7, July 2004, pp. 301-306.
- Evans, Vyvyan. 'Lexical Concepts, Cognitive Models and Meaning-Construction', *Cognitive Linguistics* 17(4), 2006, pp. 491-534.
- Goldberg, Adele E. *Constructions at Work: The Nature of Generalization in Language*. Oxford: Oxford University Press, 2006.
- Goldberg, Adele E. *Constructions: A Construction Grammar Approach to Argument Structure*. Chicago: The University of Chicago Press, 1995.
- Golder, Scott and Bernardo A. Huberman. 'The Structure of Collaborative Tagging Systems', *Journal of Information Science* 32(2) (2006): 198-208.
- Guy, Marieke and Emma Tonkin. 'Folksonomies: Tidying Up Tags?', *D-Lib Magazine* 12 (1) (2006), <http://www.dlib.org/dlib/january06/guy/01guy.html>.
- Kress, Gunther. *Literacy in the New Media Age*. London and New York: Routledge, 2003.
- Lakoff, Georg and Mark Johnson. *Metaphors We Live By*. Chicago: Chicago University Press, 1980.
- Langacker, Ronald W. *Cognitive Grammar: A Basic Introduction*. Oxford: Oxford University Press, 2008;
- Langacker, Ronald W. *Foundations of Cognitive Grammar*. Vol. 2: *Descriptive Applications*. Stanford, Cal.: Stanford University Press, 1991.
- Lévy, Pierre. *Collective Intelligence: Mankind's Emerging World in Cyberspace*. Cambridge, Ma.: Perseus Books, 1997.
- Marlow, Cameron, Mor Namaan, Danah Boyd, and Marc Davis. 'HT06, Tagging Paper, Taxonomy, Flickr, Academic Article, ToRead', *Proceedings of the Seventeenth Conference on Hypertext and Hypermedia*. New York: Association for Computing Machinery (ACM), 2006. Available at <http://www.danah.org/papers/Hypertext2006.pdf>.
- Mathes, Adam. 'Folksonomies – Cooperative Classification and Communication Through Metadata' (2004), [http://blog.namics.com/2005/Folksonomies\\_Cooperative\\_Classification.pdf](http://blog.namics.com/2005/Folksonomies_Cooperative_Classification.pdf).
- Merholz, Peter. 'Metadata for the Masses,' (2004) <http://www.adaptivepath.com/ideas/essays/archives/000361print.php>.
- Metz, Christian. 'Cinéma: Langue ou Langage?', in Christian Metz. *Essais sur la Signification au Cinéma*. Tome 1. Paris: Klincksieck, 1983, pp. 39-93.
- Mika, Peter. 'Ontologies Are Us: A Unified Model of Social Networks and Semantics,' *Journal of Web Semantics: Sciences, Services and Agents on the World Wide Web* 5(1) (2007): 5-15.
- O'Reilly, Tim. 'What is Web 2.0: Design Patterns and Business Models For The Next Generation of Software.' (2005) <http://www.oreilynet.com/lpt/a/6228>.
- Shirky, Clay. 'Ontology is Overrated: Categories, Links, and Tags', *Clay Shirky's Writings About the Internet: Economics & Culture, Media & Community, Open Source* (2005) [http://shirky.com/writings/ontology\\_overnated.html](http://shirky.com/writings/ontology_overnated.html).
- Siewierska, Anna. *Functional Grammar*, London and New York: Routledge, 1991.
- Simons, Jan. 'Tag-elese or The Language of Tags,' *Fibreculture Journal*, Issue 12, 'Models, Metamodels and Contemporary Media,' (2008) [http://journal.fibreculture.org/issue12/issue12\\_simons.html](http://journal.fibreculture.org/issue12/issue12_simons.html).
- Tonkin, Emma. 'Between Symbol and Language-In-Use', Doan, Bich-Liên, Jose Joemon and Massimo Melucci (eds.) *Proceedings of the 2<sup>nd</sup> International Workshop on Context-Based Information Retrieval*. Roskilde, Dk: Roskilde University, 2007, pp. 113-119.
- Vander Wal, Thomas. 'Explaining and Showing Broad and Narrow Folksonomies', (2005b) [http://www.personalincloud.com/2005/02/expaining\\_and\\_.html](http://www.personalincloud.com/2005/02/expaining_and_.html).
- Walker, Jill. 'Feral Hypertext: When Hypertext Literature Escapes Control', *Proceedings of the*

*Sixteenth ACM Conference on Hypertext and Hypermedia, Hypertext '05*. New York: ACM Press, 2005.

Wittgenstein, Ludwig. *Tractatus Logico Philosophicus*. London and New York: Routledge, 2001 [1922].

Zacks, Jeffrey M. and Barbara Tversky. 'Event Structure in Perception and Conception', *Psychological Bulletin* 127(1), 2001, pp. 3-21.