"thinkflickrthink": a case study on strategic tagging.

Eugenio Tisselli

The growth both in quantity and diversity of on-line communities across the World Wide Web, along with a number of new technologies that enhance both social interaction and content management, have bred an array of increasingly participatory practices. Users are engaged in bustling environments in which they can express themselves and interact with other users, creating and fostering all sorts of relationships, while uploading and sharing multimedia contents. Such environments turn into vital territories for many of their users, who can become extremely sensitive and protective of what they believe to be their rights. Thus even a small, unfavorable change in the structure of the site or in its usage policies can trigger discontent and active opposition. Actions performed by the site administrators, such as the deletion of content or the suspension of user accounts, can be perceived as abusive by the community and trigger outrage. In such situations, many uncoordinated forms of spontaneous protest and defense can emerge from the network of users. The creativity and effectiveness of these initiatives can vary greatly, with protests ranging from discussions on forums and blogs, to site-blocking boycotts[1]. This research analyzes one particular protest strategy recently adopted by a number of users of Flickr[2], a popular image-sharing site: the use of anti-censorship tags to make the protest visible within the site itself.

About tagging

Many Web 2.0 sites such as Flickr, del.icio.us[3] or Last.fm[4], offer their users the possibility of tagging online content. Tagging can be defined as the enrichment of digital contents with semantically meaningful information in the form of freely chosen text labels, or tags [1]. The freedom implied in this activity comes from the fact that tagging does not rely on a controlled vocabulary or a predefined taxonomic structure [2], but is instead an essentially individual act of classification. Tagging is fundamentally about sense-making, and can be viewed as user-defined filtering [4]. A reason that may explain why tagging has become such a popular way of on-line classification is its simplicity: a tagger must only select or upload content to a centralized database, and assign words (tags) to this material. Even though tagging can be considered mainly as an individual activity, the aggregation of tags produced by an online community evolves into a common vocabulary known as folksonomy[5]. Cattuto et al. have noted that the emergence of a folksonomy exhibits aspects also observed in human languages, such as the crystallization of naming conventions, competitions between terms or the appearance of widely spread neologisms [1]. Steels [6] sees tagging as an example of distributed cognition, based on the argument that language should be viewed "as a complex adaptive system in which a distributed group of agents collectively invent and align shared symbol systems". Steels also points out that tags, in a web context, act more like future aids for navigation in large information sources, rather than being referential markers for

---

[1] See for example: "DVD DRM row sparks user rebellion"
http://news.bbc.co.uk/2/hi/technology/6615047.stm
[2] http://www.flickr.com
[3] http://del.icio.us
[4] http://last.fm
[5] http://www.vanderwal.net/folksonomy.html

describing and discriminating objects. Tagging, in general, can be described as a highly subjective naming activity.

Golder and Huberman [2] propose a general classification of tags by their function, organized by Zollers [8] in the following table:

Tag functions	Possible motivations
Identifying what or who	organizational, attract attention
Identifying what it is	organizational
Identifying who owns it	organizational, contribution and sharing
Refining categories	organizational, play and competition
Identifying qualities or characteristics	organizational, express opinion, play and competition
Self-reference	organizational, self-presentation
Task organizing	organizational

Table 1. Classification of tags according to their function.

In Flickr, tagging is not mandatory, and unlike other resource sharing sites, deals exclusively with user-generated content, which can only be tagged by its owner[6]. Following the tag support typology proposed by Marlow et al. [4], the tagging mode supported by Flickr is "viewable", implying that users can view each other's tags and thus share ways of naming. This mode can lead to the convergence of local folksonomies, or alignment of vocabularies. There is an expectation that such localized convergences can result in global semantic effects if large numbers of users are involved. Ontologies, thus, can become an emergent feature of the community, as opposed to a pre-fixed contract or "dictionary" [5]. Flickr shows this emergent dictionary in a format that has become quite popular along with tagging: the "tag cloud", a list of tags highlighted in a way that is proportional to different criteria, such as usage frequency[7].

While this perspective appears quite promising, it presents some major and yet unsolved problems, such as synonymy, indistinct use of plurals or parts of speech or even conflicting morphological constructions and the use of upper or lower case letters or special characters to denote words. Because of these current limitations, there is a great probability that many "desired" items will never be retrieved by a query simply because they were tagged in ways which are different than expected. Moreover, the unstructured nature of tags makes it very hard to arrange them into categories, which could greatly enhance navigation through the tagged information.

Tagging as a strategy for protest

This research studies the reactions of the users of Flickr who, in an uncoordinated way, responded to a change in its filtering policy in Germany. In particular, it focuses on the birth and dissemination of a new anti-censorship tag created for the occasion: "thinkflickrthink". This event presented a unique opportunity to analyze, from its origins, the semiotic dynamics of a specific word within a social network. The creation of this and other competing anti-censorship tags points towards a strategic, even subversive use of the system itself. Users swarmed Flickr with the newly created tags, aiming at making them popular and thus visible within the "tag cloud", and also at other prominent locations throughout the site.

On June 12th 2007, Flickr announced new restrictions for its German users, which would prevent them from viewing content marked (by Flickr users themselves) as inappropriate,

---

[6] Users can also allow "friends" (contacts) to tag their photos. However, this mode of tagging was not observed during this research.
[7] http://www.flickr.com/photos/tags/

or "unsafe". The first reactions in the Flickr forum and other Internet-based media were a mix of amazement and rage[8]. Very quickly, users started to protest in different ways, one of which was the creation and usage of various anti-censorship tags. The tag "thinkflickrthink" rapidly stood out against its competing tags. On June 13th, 381 photos were tagged with "thinkflickrthink"; the next day, this number had increased by 953. On the 14th, 157 different protesters used this tag, and 163 did it the next day. By June 17th, the total number of photos tagged as "thinkflickrthink" was of 2183. This activity peaked in June 20th, with 1854 photos tagged by 60 users on a single day. Successive peaks, the first of which is largely due to massive tagging made by a single user, came in days June 27th, with 1535 tag assignments, and July 8th, with 330. Since that date protest tagging has continued, although it has significantly slowed down, with a stable average just below 100 tag assignments per day. The total number of photos tagged as "thinkflickrthink", measured on July 30th, is of 7.815. These images were tagged by 1.339 different users[9].

Figure 1 shows the tagging pattern during the period starting on June 13th (the first appearance of "thinkflickrthink") and ending on July 30th.
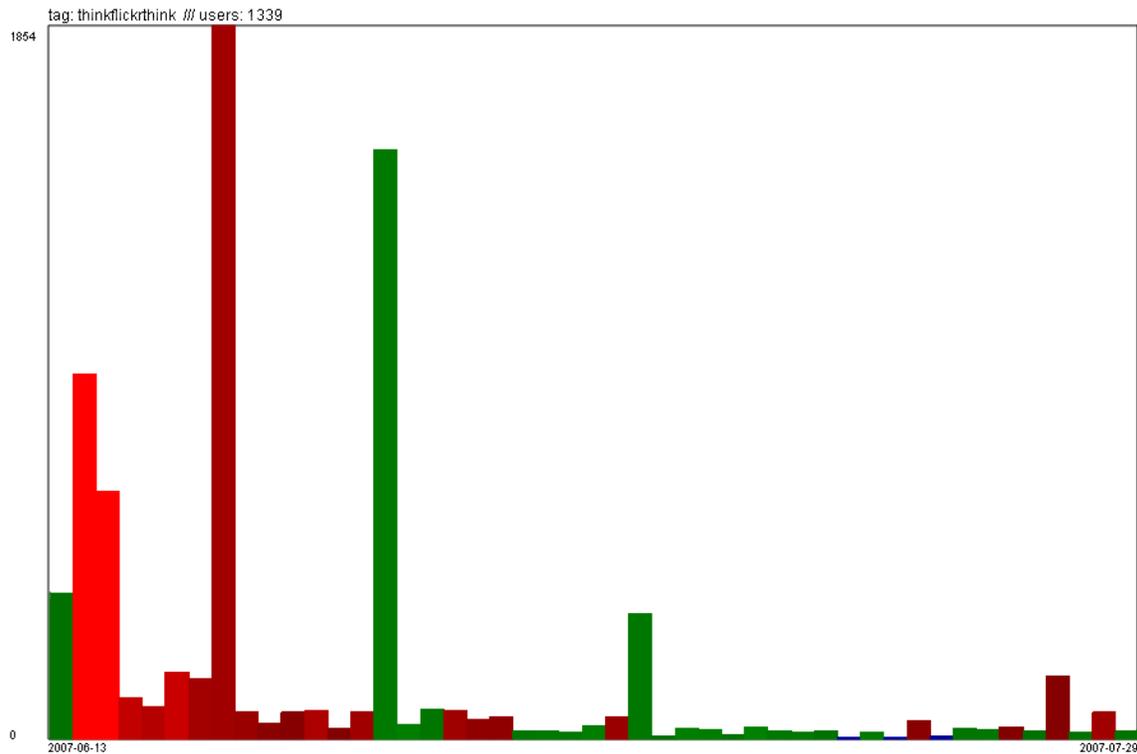

Figure 1. Timeline of images with the tag "thinkflickrthink". X-axis represents time; Y-axis represents the number of tagged images.

In Figure 1, the days in which many protesters (above 50) used the tag are shown as red bars, each with a brightness level proportional to the number of users. Green bars are used for days in which a number between 30 and 50 protesters were involved, while blue bars show days with less than 10 active users.

Figure 2 shows the growth of the tag "thinkflickrthink" over time, comparing it to other two competing tags that were also created and used for the occasion: "censr" and "againstflickrcensorship".
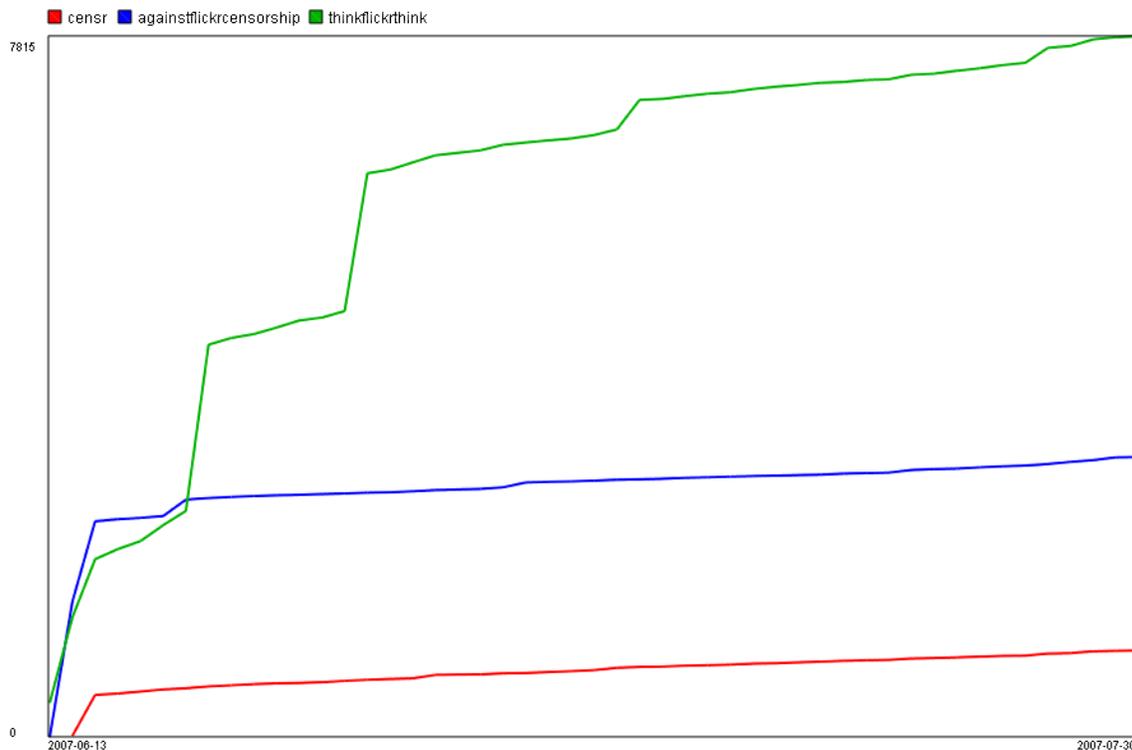

Figure 2. Timeline of the evolution of three competing anti-censorship tags. X-axis represents time; Y-axis represents number of tag occurrences.

All three tags were created between June 13th and 15th. Although all of them are still being used (along with other minor anti-censorship tags), "thinkflickrthink" has clearly become the most popular one. It is interesting to see that, on July 14th, the tag "againstflickrcensorship" had actually been used in more posts (1.511) than "thinkflickrthink" (1.334). However, it had been used by fewer users (26 against 170). In many images, several anti-censorship tags were used simultaneously: a total of 16 different explicit anti-censorship tags were detected. The top co-occurring tags with "thinkflickrthink" were "censorship" (1457 co-occurrences), "zensur" (899) and "protest" (818).

An interesting fact that arises from the tag dynamics analysis is that not all the tagged photos were uploaded at the same moment they were used to protest. Most of them had been uploaded previously, and only had the anti-censorship tags added to them after the critical date, June 13th. The graph shown Figure 1 is based on the dates in which individual photos were modified; this means that the time-scale reflects the dates in which a change was made to the images, such as the association of a new tag. If we look instead at the dates in which the photos tagged with "thinkflickrthink" were uploaded, we will find that, from a total of 7.815 photos, 5.701 (72,94%) were added to Flickr before June 13th. This clearly points towards a strategic use of tagging: not for describing, but for quickly swarming the system and turning the newly invented tag into a popular (and

thus visible) one. Users took many of their previous images and just added the new tag to them.

Figure 3 shows the upload pattern of the photos tagged "thinkflickrthink". It appears that only relatively recent photos were re-tagged, with the "oldest" one being from January 12th, 2007. This figure uses the same color scheme as Figure 1.
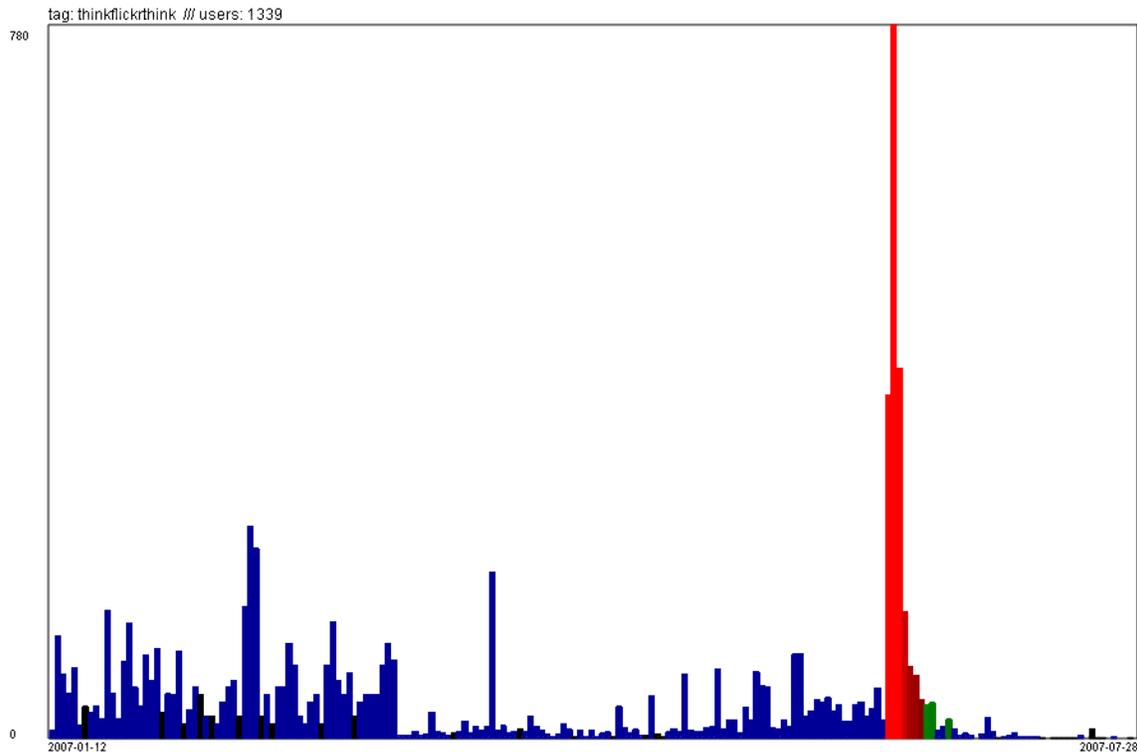

Figure 3. Timeline showing the number of images tagged with "thinkflickrthink", according to their upload date. X-axis represents time; Y-axis represents the number of uploaded images.

In Figure 3, the maximum peak happens also on June 14th, with 654 users uploading 780 photos on a single day. If we browse through Flickr searching for images with the protest tag, we will find pictures created after the beginning of the protests that explicitly make a visual anti-censorship statement, along with many others whose visual content holds no relation to the tag.

The posting patterns shown in Figure 1 and Figure 3 are practically the same after June 27th, suggesting that the strategy of re-tagging old photos was only used during the first couple of weeks of the protests. 34 different protesters (2,53%) who applied the re-tagging strategy were found.

The network of protesters
The high speed at which the tag "thinkflickrthink" spread through Flickr immediately suggests the existence a strongly-connected network of users. In order to study how the protest tag users were interconnected within Flickr, the public contact list of every user was obtained. Since one of the most popular features at Flickr is the possibility of forming groups of users with common interests, a query for obtaining the anti-censorship groups to which these users belonged was also made.

During the anti-censorship protests, users got together into newly-formed protest groups. In particular, the group "Against Censorship at Flickr!" experienced an impressive growth: in the first week of the anti-censorship campaign, its number of members went from 0 to almost 12.000. Besides this major group, other 11 anti-censorship groups were detected.

The users of "thinkflickrthink", by their mere usage of the tag, can be considered as a group, albeit an emergent one, in which its "members" share the common use of a strategy. Of all the involved users, 1.115 of them (83,27%) have at least one contact within this spontaneous group. The number reflects the total users who "know" or "are known by" another user, regardless of reciprocity. Contacts in Flickr are unidirectional: user A is in user B's list, but this doesn't imply that B is in A's. The fact that the majority of users are interconnected suggests that the new tag may have spread mainly through contact with "close neighbors", although the anti-censorship groups played an important role as well.

The following list offers a quick view of the contact distribution within the group of protesters:

- Users with more than 30 outgoing contacts: 6 (0,44%)
- Users with 10 to 20 outgoing contacts: 89 (6,64%)
- Users with less than 10 outgoing contacts: 867 (64,00%)
- Users with only incoming contacts: 163 (12,17%)
- Non-connected users: 224 (16,72%)

Contrarily to what may be expected, there is very little reciprocity in this network. The number of bidirectional relationships, (pairs of protesters who are both on each others' contact lists) is quite low: only 46 of the 4.320 (1,06%) contacts are bidirectional.

The majority (76,23%) of the 1.115 interconnected users belong to at least one anti-censorship group. Respectively, 63,83% of the 244 non-interconnected users were also members of an anti-censorship group. By taking the parameters "interconnectedness" and "affiliation to an anti-censorship group" into account, the following classification of protesters can be proposed:

Group	Quantity	Characterization	Total posts
Activists	850 (63,48%)	users who are interconnected and belong to an anti-censorship group	6.848 (87,62%)
Friends	265 (19,79%)	users who are interconnected but don't belong to an anti-censorship group	372 (4,76%)
Newcomers	143 (10,67%)	users who are not interconnected and belong to an anti-censorship group	482 (6,16%)
Supportive bystanders	81 (6,04%)	users who are not interconnected and don't belong to any anti-censorship group	113 (1,44%)

Table 2. Characterization of "thinkflickrthink" users.

The "Activists" are not only the majority in this group, but also the most active. Nevertheless, their activity is a highly unequal one: 72.98% of the posts were made by only 5 users (0,58%), while 621 users (73,05%) made exactly one post (9,06%).

The posting pattern of "Friends" is much more regular than that of the "Activists": 25 users (9,43%) made 132 posts (35,48%), while 240 of them (90,56%) posted one image each (65,51%). The users in this group can be considered as "one-time" posters.

"Newcomers" also present inequalities: 1 user (0,69%) made 306 postings (63,48%). The number of users with one post each is 119 (83,21%); this amounts to 24,68% of the total posts by this group. "Newcomers" also seem to be "one-timers", with only an outstanding case of hyperactivity.

"Supportive bystanders" can be considered as "one-timers" too, with 90,12% of them (73 users) having posted 64,60% of the images by making one post each. They show no extreme cases.

With "Activists" and "Friends" as a majority, it is possible to think that "knowing" someone who has used the protest tag is a more decisive factor for engaging into strategic tagging than that of belonging to an anti-protest group. The contact list could be considered then as the primary source of consensus. In order to verify the validity of this hypothesis, the average clustering coefficient of the social network of connected protesters was calculated [6]. The clustering coefficient is a structural measure that indicates the presence of groups (or clusters) within the neighborhood of a particular node in a social network. The network was modeled as a directed graph, with protesters as nodes and contacts as edges. The average clustering coefficient C of the graph is 0,0944 (1.115 nodes, 4,53 average degree (edges) per node). This means that, on average, ~9% of any user's contacts will have links between them. While this value can not be considered as being high by itself, it is significantly higher than the average clustering coefficient (0,0044) of a randomly generated graph with the same topological parameters. According to Watts and Strogatz, this relation suggests a network structure that can be potentially characterized as being a "small world". "Small world" networks have been found to be highly efficient structures for the propagation of information [3]. Thus, "thinkflickrthink" may have benefited from this particular topological feature of the network of protesters, by quickly spreading through their contact lists.

It is important to observe the presence of unconnected protesters in the group. Even if they may seem isolated or detached, the fact that they are located at the group's borders makes them potential disseminators of the protest and thereby of "thinkflickrthink" to other "remote" sub-networks formed by their own contact lists.

Power users

A deeper analysis of the most active users shows that, out of 1.339 users, only 6 of them (0,44%) made the 67,68% of the total posts. All of these users were categorized as "Activists", except one of them, who is a "Newcomer". We found that this "Newcomer" is not a German user, which shows that the intensity of the protest was not limited to a geographical location. We can think of "Newcomers" also as "Foreign supporters".

All of these hyperactive protesters used the strategic re-tagging of old images, and produced only limited outbursts of activity. Most of their images were re-tagged at once. Remarkably, none of them belongs to the group of the "most connected" users, who are all "Activists". Despite their impact on the overall results, they are not vital for the existence of the network itself, since they are not connectors of a significant number of users.

The roles of protesters

In addition to this taxonomy of users, we also propose a different overlapping categorization, which describes their roles in the protest. These roles are "Strategists", or users who re-tagged their old images, "Initiators", users who were among the first ones to use the "thinkflickrthink"[10], and "Followers", who are those users that are neither "Strategists" nor "Initiators". We present the distribution of users classified by role in Table 6.

---

[10] This category includes protestors who used the tag on the first three days, and excludes "Strategists".

Role	Number of users	Type of protestors
Strategist	34 (2,53%)	Activists: 24 (70,58%) Newcomers: 2 (5,88%) Friends: 7 (20,58%) Supportive bystanders: 1 (2,94%)
Initiator	162 (12,09%)	Activists: 80 (49,38%) Newcomers: 14 (8,64%) Friends: 48 (29,62%) Supportive bystanders: 20 (12,34%)
Strategist and Initiator	6 (0,44%)	Activists: 3 (50,00%) Newcomers: 0 Friends: 2 (33,33%) Supportive bystanders (16,66%)
Follower	1.143 (87,45%)	Activists: 746 (65,26%) Newcomers: 127 (11,11%) Friends: 210 (18,37%) Supportive bystanders: 60 (5,24%)

Table 3. The roles of protesters.

Conclusions

A tag can be created and disseminated for strategic purposes. Thus, the classification of tags proposed by Golder and Huberman [2] has to be extended to include this usage, which can be characterized as being metalinguistic. Zollers has identified what she calls "Activism" in the use of tagging [8], however, the term "strategic" is preferred here, since it presents broader connotations. A strategy is a long term plan of action designed to achieve a particular goal. By striving for visibility, strategic taggers went beyond mere description, annotation or even expression, and tried to subvert the system by exploiting its own features. They tried to expand the limits of the linguistic context of tagging, in order to be able to speak loudly and directly to those who run the website.

The analysis of the data shows that protestors most likely disseminated the use of strategic tagging among their contacts, rather than within a particular specific-interest group. A list of contacts is much closer to a hand-picked ensemble of friends than one of such groups, and therefore represents a bigger influence for the list's owner.

The categorizations of users proposed can be practical tools to be applied in other similar cases, and used to analyze the different patterns of activity of subgroups that spontaneously emerge from networks. The fact that most of the collective protest tagging was done by a minority of protestors reflects a phenomenon that should be acknowledged when dealing with communities: the actions of a few can outweigh those of the many. Collectivity, in this case, doesn't necessarily stand for proportionality, at least in quantitative terms.

The study of the dynamics of uncoordinated semantic strategies within dense on-line communities is of enormous importance to gain a greater understanding of how social and linguistic interaction takes place in a technological environment, and how it can augment the users' potential for direct action.

References

[1] C. Cattuto, V. Loreto, L. Pietronero, Collaborative Tagging and Semiotic Dynamics. PNAS, 104(5):1461--1464, 2006.

[2] S. Golder, B. Huberman, The Structure of Collaborative Tagging Systems. Information Dynamics Lab, HP Labs. (http://www.hpl.hp.com/research/idl/papers/tags/tags.pdf)

[3] J. Kleinberg, The Small World Phenomenon: An Algorithmic Perspective. Proceedings of the thirty-second annual ACM symposium on Theory of computing. Portland, United States, 2000.

[4] C. Marlow, M. Naaman, D. Boyd, M. Davis, Position Paper, Tagging, Taxonomy, Flickr, Article, ToRead. Collaborative Web Tagging Workshop at WWW 2006. Edinburgh, Scotland, May 22, 2006.

[5] P. Mika, Ontologies Are Us: A Unified Model of Social Networks and Semantics. Journal of Web Semantics, Elsevier, Volume 5, Number 1, p.5--15 (2007)

[6] L. Steels, Collaborative Tagging as Distributed Cognition. Pragmatics and Cognition, 14(2):275--285, 2006.

[7] D. J. Watts, S. H. Strogatz, Collective dynamics of 'small-world' networks. Nature vol. 393, 4 June 1998

[8] A. Zollers, Emerging Motivations for Tagging: Expression, Performance and Activism. WWW2007, Banff, Canada, May 8–12, 2007