From the crowd to communities: New interfaces for social tagging

Chao Zhou and Aurélien Bénel

Laboratoire ICD/Tech-CICO (FRE CNRS 2848)
Université de technologie de Troyes
12 rue Marie Curie, BP 2060, 10010 Troyes Cedex, France
{chao.zhou, aurelien.benel}@utt.fr

Abstract. Social tagging is intimately linked to 'tag cloud', the visualization apparatus which is intended to bring the 'wisdom of crowds'. But what is suited for the 'crowd' may be not for communities. In this article, we propose a new interface for social tagging in collaborative systems that includes several improvements: multi-viewpoints, multi-tags selection, and tags relations. We illustrate this apparatus on the collaborative analysis of a scientific archive.

Keywords: Visual Interfaces, Web 2.0, Social tagging, Scientific archives, Qualitative analysis.

1 Introduction

Social tagging is a recent practice in which every user of a computer system shares the free keywords she use to categorize document resources. It is intimately linked with a visualization apparatus called 'tag cloud' in which these keywords are usually displayed in alphabetical order and visually weighted by font size depending on the number of people who used it. This visual interface efficiently provides a preattentive overall visualization of the trends, sometimes seen as the 'wisdom of crowds' [1]. However what is suited for the 'crowd' may be not for communities of interest.

In this paper, we introduce a set of different implementations of tag clouds that are widely used in web 2.0 applications. We describe our investigation into why the current interface for social tagging (tag cloud) works for crowd but will not for communities. In the last section, we propose a new interface for social tagging in collaborative systems. In the meantime, we illustrate it on the collaborative analysis of a scientific archive.

2 Visualizing Participative Tagging

The first web site to use intensively a tag cloud was *Flickr*, but the idea of *Flickr*'s tag clouds was likely inspired by a blog plug-in called *Zeitgeist* (Fig. 1) created by Jim Flanagan in 2002.

2.1 Zeitgeist

'Zeitgeist' is a German expression that means "the spirit (Geist) of the time (Zeit)". It denotes the general intellectual, moral, and cultural climate of an era.

Zeitgeist was a blog plug-in which was designed to read web server referrer logs and parse the referrals from search engines to get the search terms which led people to the blog [2]. Actually, Zeitgeist is not a tag cloud but a weighted list of search engine queries, in which font size is correlated with their popularity. Nevertheless, Zeitgeist already has the global appearance of a tag cloud.

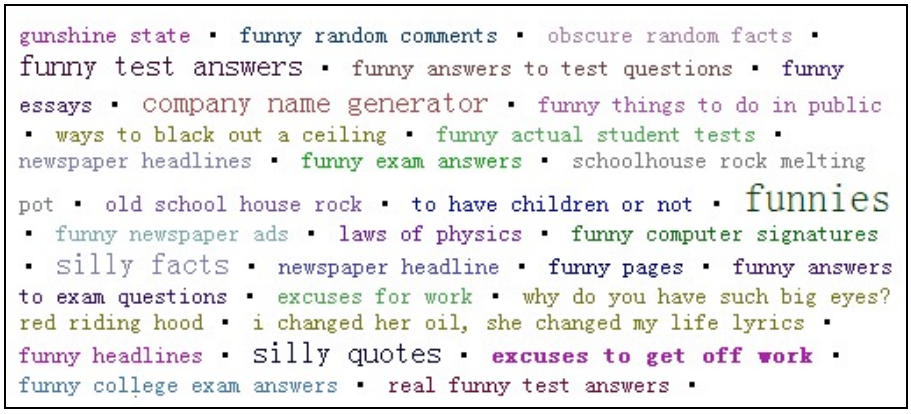

Fig. 1. An Example of Jim Flanagan's Referral Zeitgeist

2.2 Flickr

Flickr is an online community platform where users can share their personal photographs. It is an early example of web 2.0 applications. *Flickr* was one of the first websites to implement a tag cloud. *Flickr* users can use tag cloud to browse and re-find their photographs. *Flickr*'s tag cloud had made some improvements from *Zeitgeist*: single word tags from community instead of search engine phrases, alphabetical word order, a single color and an attractive font, etc.

Furthermore, *Flickr* provides several tag clouds implementation on the same web page (Fig. 2) to present different levels of popularity decay, respectively "in the last 24 hours", "over the last week" and "all the time most popular tags".

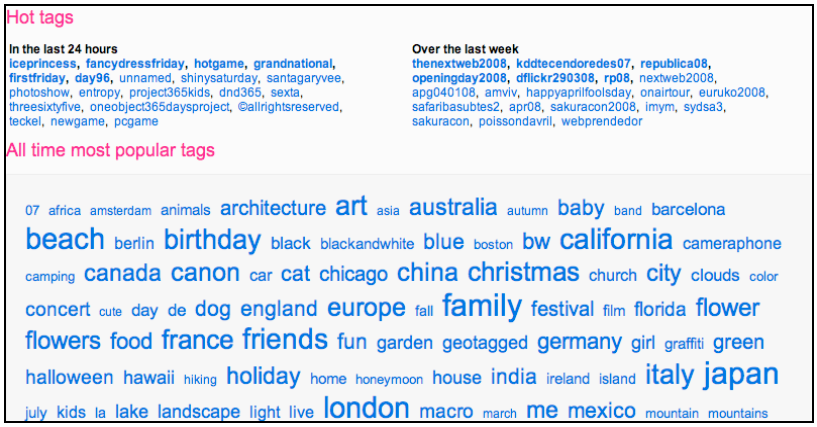

Fig. 2. *Flickr*'s popular tags

2.3 Del.icio.us

Del.icio.us provides a social bookmarking service, the users of *Del.icio.us* can use tags to organize, share and discover bookmarks online. The main difference between *Flickr* and *Del.icio.us* is that *Flickr*'s tags are mainly used by users to manage their own photographs, on the contrary the tags of *Del.icio.us* are widely used to describe the content that were written by other community members. *Del.icio.us* users can attach a tag or several tags to a link when they are saving a bookmark. From a user's personal page (Fig. 3) we can see the time when the links were saved, the number of people who also had saved them, and the tags that other people had attached to them. Users can also group related tags into a bundle, and a tag can belong to different bundles at the same time. Bundle is a handy feature, especially for people who have vast number of tags.

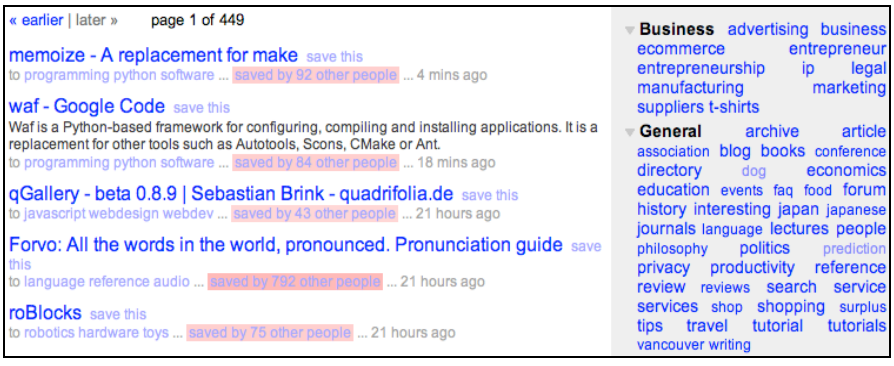

Fig. 3. A personal page on *Del.icio.us*

Del.icio.us is considered as a social web service, not only a user can see his own bookmarks, but also he can see other users' bookmarks and how they tag it. *Del.icio.us* has a tag cloud (Fig. 4) of popular tags on its home page. The tag cloud of popular tags could be ordered either alphabetically or by popularity. Browsing other users' tags or popular tags helps users find interesting content.

Fig. 4. The popular tags on *Del.icio.us*

2.4 Diverted Uses of Tag Clouds

It is worth noticing that tag clouds are more and more used in other domains than participative tagging such as mind mapping of graphic design (Fig. 5). Their diverted uses could be seen as a hint of their aesthetic and semiotic power.

Fig. 5. "Web 2.0 map" (created by Markus Angermeier & Luca Cremonini)

Bertin's semiotic features [3] could explain the efficiency of tag clouds representations. Whereas tags themselves require a time-consuming process to be read, some important tag features can be got at a glance (Preattentive processing). Viewers can easily know tags trend because of the size semiotic feature, which can express a proportion. They can also locate a tag without any attention because of the

difference of size in adjacent tags. On the contrary, if tags are ordered by popularity, tags will lose their unique visual property.

3 From Participation to Collaboration

The tag clouds we studied earlier seem to be very well suited to "participatory tagging" which aims at re-finding their own tagged resources and discovering tagging trends. However, it does not match the needs of "collaborative tagging" as it could be done in a community. For example, scientific communities and corporate teams have to cope with interpretation conflicts and consensus building, and this brings a new level of complexity.

3.1. Interpretation Conflicts

"Diversity of opinion, independence, decentralization, and aggregation" are said [1] to be the main characteristics of participative settings. Because only the trends count in a tag cloud, the diversity of opinions, needs and languages [4, 5] can be aggregated in a single representation. However this aggregation of personal collections does not make a collective collection [6].

Indeed the needs are quite different in collaborative settings. Because members have to take into account each other's viewpoint, conflicts are unavoidable. It is worth noting that in Wikipedia for example, the only way to write collaboratively articles from a "neutral point of view" is to provide discussion pages and revision history on which every edit is authored and a dated [7]. More generally, the need for providing intersubjectivity by distinguishing and comparing viewpoints can be seen as the main feature of an "interpretation assistance system" [8].

3.2. Consensus Building

Once interpretation conflicts permit to distinguish different viewpoints, it is then possible for different people to choose the same viewpoint and co-create it. But sharing a mere word list is not building a consensus. The usual way to build a consensus is trying to organize terms collectively (*e.g.* phylogenetic taxonomies in biology, vases typologies in archaeology, UML models in information technology…). One should note that making tacit models explicit in search of consensus often leads in fact to the creation of new viewpoints.

An interesting thing is that some folksonomy users appeared to be attempting to establish a hierarchical structure by building up a "pathway" within the tag [6]. For example, on del.icio.us users tagged some web resources on the subject of programming language with the tag 'programming/c++', 'devel/java', 'webdesign/css'. However, folksonomies are pure combinatory (*vs* classificatory) indexing systems, which means that they do not allow creating relationships between tags.

4 Visualizing Collaborative Tagging

This section presents improvements on tag clouds for their use by communities. To illustrate them, we upgraded *Porphyry*[1], our software prototype, and then we applied it to the studies made by a research team[2] in history of art and archaeology. Those studies deal with the iconography of Dionysos and banquets on vases from the area of Paestum (Italy). To do this, the team gathered more than 600 photographs about those vases from museums all over the world. Three master students[3] used *Porphyry* to analyze a subset of the vases showing an altar (called "bômos" in greek) (Fig. 6).

Contrary to the original graph-based visualization used in Porphyry, the new cloud-based visualization makes it possible to get an overview of the description of the corpus at first sight, without having to scroll. Because the analysis is collaborative and not participative, the font size does not depend on the number of uses by different people but only on the number of uses on different documents.

Fig. 6. Studies about altar representations on vases (*Porphyry* screenshot)

4.1. Multiple Tag Clouds

As we said earlier, we must distinguish viewpoints from different people or teams in order to show interpretation conflicts. In our novel interface, it is possible to load

[1] http://www.hypertopic.org/index.php/Porphyry

[2] Jean-Marc Luce, Pascale Jacquet & Véronique Pouyadou, CRATA Laboratory, University of Toulouse II, France.

[3] Marion Lagarde, Elodie Lacrampe, Clélia Robinet.

several viewpoints at the same time with every viewpoint represented by its own tag cloud. In our example (Fig. 7), the master students have been able to 'interweave' their study with a former postdoctoral study about the vases themselves (form, supposed date and painter).

We adapted a filtered browsing algorithm (formerly developed for graphs [9]) to tag clouds. Clicking on a tag such as 'coupe sans pied' (non-stemmed cup) triggers the computation of not only the documents results but also of every tag cloud to reflect the new document corpus. Seeing the results of tags selection from a viewpoint into another provides a powerful way to compare viewpoints on a same corpus.

The main problem was that documents items levels could be different: the photograph of a vase side and the folder containing the photographs of a given vase. Therefore, we introduced a tabbed pane system in which it is possible to focus on a document item level, and in which if a document part is dealing with a tag, the whole document is considered to do so.

Fig. 7. Interweaving two studies on the same corpus (*Porphyry* screenshot)

4.2. Tags Relations

So that a team can co-create a consensus, we decided to allow them to organize tags with relations. However, showing arrows on a tag cloud is not straightforward. Most of graph drawing algorithms impose a given layout for nodes, incompatible with the cloud aspect. We could have used a 'semantic' ordering of tags [10, 11] to reduce

links overlapping, but the whole link structure is too complex for a preattentive perception anyway. Therefore we prefer an interactive partial visualization.

For example (Fig. 8), when the mouse moves over the tag '*récipent*' ('container' in French), it shows the arrows from '*récipent*' to the parent tag 'objet' (object) and to the child tags 'ciste' (a kind of basket) and '*vase*'. If the user follows the arrow and moves the mouse over '*vase*', it shows the arrows to child tags like '*amphore*' (amphora), '*outre*' (goatskin), or '*phiale*' (a libation vessel).

Fig. 8. When the mouse is over a tag, the links to other tags are shown (Porphyry screenshot)

5. Conclusion

In this paper, we first introduced tag clouds with the precursory *Zeitgeist*, the popular *Flickr* and *Del.icio.us*, and the latest diverted uses of tag clouds in graphic design. We analyzed then why tag clouds, which are so efficient in participatory settings, do not fit the needs of a collaborative setting, in particular because of their inability to support interpretation conflicts and consensus building. Finally we proposed to improve tag clouds with multi-viewpoints and tag links and we illustrate their use on a scientific archive in archaeology.

We currently try to complement the intersubjective visualization we presented in this paper with a diachronic visualization. The preattentive perception of a whole tagging history is although still a challenge.

References

1. Surowiecki, J.: The Wisdom of Crowds: Why the Many Are Smarter Than the Few and How Collective Wisdom Shapes Business, Economies, Societies and Nations. Doubleday, New York (2004).
2. Flanagan, J.: Zeitcode Module Documentation. Wiki page (2005), http://wiki.dlma.com/doku.php?id=zeitcode:zeitcode_doc
3. Bertin J.: La graphique. Seuil, Paris (1970)

4. Mathes, A.: Folksonomies - Cooperative Classification and Communication through Shared Metadata. Blog post (2004), http://www.adammathes.com/academic/computer-mediated-communication/folksonomies.html
5. Quintarelli, E.: Folksonomies: Power to the People. In: 1st ISKO Italian Chapter (2005). http://www.iskoi.org/doc/folksonomies.htm
6. Guy, M., Tonkin, E.: Folksonomies: Tidying up Tags? D-Lib Magazine, vol. 12 (2006). http://www.dlib.org/dlib/january06/guy/01guy.html
7. Viègas, F. B., Wattenberg, M., Dave, K.: Studying Cooperation and Conflict between Authors with History Flow Visualizations. In: CHI 2004. pp.575--582. (2004) http://web.media.mit.edu/~fviegas/papers/history_flow.pdf
8. Bénel, A., Egyed-Zsigmond, E., Prié, Y., Calabretto, S., Mille, A., Iacovella, A.,Pinon, J.M.: Truth in the Digital Library: From Ontological to Hermeneutical Systems. In: ECDL 2001. LNCS, vol. 2163, pp. 366--377. Springer, Heidelberg (2001). http://benel.tech-cico.fr/publi/benel_ECDL_01.pdf
9. Bénel, A., Calabretto, S., Iacovella, A., Pinon, J.-M.: Porphyry 2001: Semantics for Scholarly Publications Retrieval. In: ISMIS 2002. LNAI, vol. 2366, pp. 351--361. Springer, Heidelberg (2002). http://benel.tech-cico.fr/publi/benel_ISMIS_02.pdf
10. Hassan-Montero, Y., Herrero-Solana, V.: Improving Tag-clouds as Visual Information Retrieval Interfaces. In: International Conference on Multidisciplinary Information Sciences and Technologies, Mrida, Spain (2006). http://www.nosolousabilidad.com/hassan/improving_tagclouds.pdf
11. Begelman, G., Keller, P., Smadja, F.: Automated tag clustering: Improving search and exploration in the tag space. In: Collaborative Web Tagging Workshop at WWW2006, Edinburgh, Scotland, Edinburgh, UK (May 22-26 2006) http://www.pui.ch/phred/automated_tag_clustering/