Human perception through collaborative semantics

A. Malizia*

Department Ingeniería Informática,
Universidad Carlos III, Madrid, Spain
E-mail: amalizia@inf.uc3m.es
*Corresponding author

A. De Angeli

Manchester Business School,
The University of Manchester Booth Street West,
Manchester M15 6PB, UK
E-mail: antonella.de-angeli@manchester.ac.uk

S. Levialdi

Dipartimento di Informatica,
Sapienza, Università di Roma, Italy
E-mail: levialdi@di.uniroma1.it

Abstract: Human perception is a crucial factor in the evaluation and use of an application or artefact. Currently, there is a relevant number of websites that rely on collaborative tagging: such systems allow users to add tags (words) for categorising contents. In this work, we explored a methodology for detecting human perception by using collaborative tagging systems. This approach highlighted the potential use of such systems evaluating the end users’ perception without requiring a significant effort in designing specific evaluation tools like, for instance, questionnaires.

Keywords: collaborative tagging systems; distributed intelligence; semantics differential; usability evaluation.


Biographical notes: Alessio Malizia is a Professor at the Dep. Ingeniería Informática, Universidad Carlos III, Madrid, Spain. He completed his PhD from Sapienza University of Rome, Italy. Since 1998, he has been researching in visualisation, human–computer interaction, and visual languages.

Antonella De Angeli is a Lecturer in Human–Computer Interaction at the Manchester Business School of the University of Manchester. Her research addresses cognitive, social and cultural consequences of information technologies. Her was awarded a PhD in Experimental Psychology from the University of Trieste (Italy). Part of her research was conducted at the OGI (Portland, USA), Loria (Nancy, France) and IRST (Trento, Italy). From 2000
A. Malizia et al.

to 2004, she worked as a Senior HCI Researcher for NCR Ltd. She has published over 80 papers, served in the programme committee of leading conferences, and reviewed for international journals.

Stefano Levialdi has worked for over 20 years in parallel image processing, moving to human–computer interaction in 1984 mainly interested in the visual channel: visual languages, visual interfaces and related usability issues. He has started the computer science curriculum at the University of Rome, Faculty of Sciences. He has published over 250 papers and started the Journal of Visual Languages and Computing (1990) and is IEEE Life Fellow. He has taught different subjects within Computer Science, his latest one is Human–Computer Interaction. He has strong international ties and cooperates with a number of foreign universities; he is fluent in Spanish, English and Italian.

1 Introduction

Collaborative design, social creativity, and meta-design are identified as themes that will be of great importance in the years to come (Fischer, 2006). Nowadays, the concept of design is used very broadly as affecting many aspects of the process of building, using, and evolving software-intensive systems. Researchers and professionals coming from different disciplines and engaging in collaborative design can contribute to social creativity by exploring new approaches. Meta-design is a methodology stimulating users to behave not only as passive individuals but also as active contributors and designers, thereby sustaining social creativity. As Gerhard Fischer stated in Fischer (2006): “the power of the unaided individual mind is highly overrated”. Although creative individuals are often thought of as working in isolation, much of our intelligence and creativity results from interaction and collaboration with other individuals (Bennis and Biederman, 1998), exploiting the ‘symmetry of ignorance’ as a source of power. Focusing on social creativity does not mean that the importance of individual creativity will be ignored (Csikszentmihalyi, 1997). Creative individuals can make a relevant difference. Individual creativity can be enhanced by supporting appropriate socio-technical settings because many creativity activities take place in social context where interactions with other people, along with artefacts allowing group knowledge, are important contributors to the overall process documents. Creativity is not only inside people’s heads, but also in the interaction between a person’s thoughts and a socio-cultural context. Situations that support social creativity need to be sufficiently open-ended and complex in that users will encounter breakdowns.

These considerations are now influencing the evolution of the web (Web 2.0) and thus the information spread across the internet. Collections of document, images and videos drive the evolution of multimedia digital libraries on the web, but all this information needs to be indexed and categorised for allowing users to search for the appropriate contents. One possible answer to the distributed knowledge needs is Collaborative Tagging Systems.
2 Collaborative tagging

Collaborative tagging is the process by which users add metadata, by means of descriptive terms (tags), to community-shared content to categorise or organise documents for future navigation, filtering or search. These tags are usually chosen informally and personally by the author/creator of a document: by contrast, in document repositories, digital libraries or other systems of content organisation, metadata creation belongs to the domain of dedicated professionals, such as librarians; or in other cases, the classification of the documents derives from additional material supplied by Mathes (2004) and Golder and Huberman (2006). Based on social structures, tagging is an adaptable process; it takes the form best supported by the content, letting users to decide the categorisation of the content, rather than imposing a rigid structure on it. As a matter of fact, collaborative tagging, allowing every user to link tags to items, is most useful in an environment like the World Wide Web, where there cannot be a single “content classification authority” and there are vast amounts of content being produced by users. Because of their recent popularity on the web, collaborative tagging systems are also an interesting alternative to the semantic web approach, as with ontologies: in case the knowledge base is built by domain-field experts in agreement with the community and found through a predetermined relationship among keywords, while in a collaborative tagging system content indexing develops as a natural process.

Nowadays, there are a number of eminent websites that rely on collaborative tagging: these systems allow users both to add tags and share content, so they can browse content categorised by other users as well as categorise information for their own purposes. Therefore, there are both personal and social aspects in the tagging activity: that is why some refer to collaborative tagging systems as ‘folksonomies’ (Mathes, 2004), short for ‘folk taxonomies’, albeit there is still some debate whether this term is accurate (Vander Wal, 2005). Examples of such sites are: Del.icio.us (http://del.icio.us), a site allowing users to bookmark, categorise and share websites, Flickr (http://www.flickr.com), a service giving the possibility to users to tag photographs they own, and Youtube (http://www.youtube.com), similar to Flickr but for videoclips.

In Vander Wal (2005), folksonomies are classified as:

- ‘broad’ folksonomies
- ‘narrow’ folksonomies.

A broad folksonomy is the result of one item being categorised by many people (Del.icio.us is an example). Such systems provide the ability to normalise on some terms. There will be some terms that are used by many people to describe one item or many items by the same terms. The concentration of terms can take advantage of power laws (like the Zipf distribution (Zipf, 1949)) to quickly see the preferred terms for an item or items. On the other hand, a narrow folksonomy is the result of one person categorising one item (Flickr is an example). It is more difficult to normalise with synonyms (or to gain acceptance) using a narrow folksonomy, nevertheless, a broader acceptance can be obtained through related tags so allowing the construction of a taxonomy.

Anyway, despite their huge proliferation over the web, there are still some disadvantages in using collaborative tagging systems. We may find a few reasons to consider a folksonomy less than ideal in many cases:
none of the current implementations have synonyms control (e.g., ‘mac’ and
‘macintosh’ do not coincide in del.icio.us)

a lack of precision involved in using simple one-word tags

no hierarchical structure

the content types are too simple.

In the following section, we focus on three collaborative tagging systems: delicious, flickr and youtube.

2.1 The real value of folksonomy: delicious, flickr and youtube

The del.icio.us website, as defined by its creator J. Schachter, is a social bookmarking
web service for storing, sharing and discovering web bookmarks; the site came online in
late 2003 and it is now part of Yahoo!. The system is established on a non-hierarchical
keyword document categorisation, where users can tag each of their bookmarks with any
number of freely chosen words. Del.icio.us is considered users’ ‘social’ since not only do
user store their own bookmark but they can also see other users’ bookmarks. A combined
view of bookmarks with a given tag is also available, as well as the most recently added
bookmarks and the most popular ones.

Moreover, information about who contributed, how many other users stored them
and, of course, the associated tags are also provided. One of the major reasons behind the
popularity of del.icio.us within the web community is that personal value comes before
network value (Spool, 2006). This can be simply explained as follows: no one can
contribute to a community with value if they have not found something valuable before.
In the case of del.icio.us, this value is represented by storing personal bookmarks:
all other utilities are subsidiary. We may consider the tagging activity as secondary,
and performed for personal benefit; even the tags’ aggregation feature, that is one of the
strong points of del.icio.us, could not exist without firstly providing users with tangible
value. Lastly, there is no doubt that exploring and understanding systems like this one is a
big step to redefine web navigation. The other two websites to be examined are other
clear examples of what we have previously said, but differently from del.icio.us
videoblogs, address multimedia content. Flickr, part of Yahoo! alike del.icio.us, is a
photo management website and an online community platform. Its real valuable feature
allows users to store and share personal pictures, furthermore the service is widely used
by bloggers as a photo repository. Youtube, recently acquired by Google, is a web service
allowing users to store and distribute their videoclips; the wide variety of site content
includes movies, tv clips and music videos, as well as amateur content such as
videoblogging.

3 Semantic differential through collaborative tagging systems

In this work, we propose a methodology for understanding the perception of an e-service
among a community of users. Our approach is based on tagging systems rather than on
traditional questionnaires. Tagging systems give users the freedom of commenting on
online content as they wish, and are very intuitive and easy to use. On the contrary,
devising effective questionnaires requires experience, time, sensibility and pilot testing. Questionnaires also require commitment for filling them in. No questionnaire can ever provide perfectly accurate measurement, as they rely on self-report data, which gather estimation rather than direct measurement. Tagging systems are more direct: users can tag a website by using the adjectives they would probably choose when communicating with, or suggesting the link to, other users. The analysis of tagging systems has the potential to detect the perception of a website shared by a group of users directly, without requiring time from the user (form filling, interviews, etc.) and also from the researchers, who need to devise the instruments and analyse the results.

Extracting meanings from collaborative tagging systems is a complex task. This is because the user can freely add tags originating from basic level variations like the same word written with uppercase and lowercase letters, synonyms and polysemy. In our previous work, we introduced a Semantic Halo technique to deal with word semantics in tagging systems (Dix et al., 2006). The basic idea consisted of using co-occurrences of tags to cluster their relationships and meanings. In Figure 1, we show an example of results obtained by our Semantic Halo algorithm for the word ‘Math’ within the del.icio.us community.

**Figure 1** The visualisation interface presented for the tag ‘Math’ (see online version for colours)

The query tag ‘Math’ is displayed in the centre of the screen and boldface. The tags belonging to node 1 are considered synonyms of the submitted tag in the folksonomy. In this example, we can see that the terms Math and mathematics are included in this set; in fact, the basic level variations, like the same terms written with a capital letter, can occur in tagging systems but they are still managed by the algorithm. Tags belonging to node 2 represent the generalisation of the submitted tag, like: reference, formal and
statistics (some interpretation error might occur). Tags belonging to node 3 represent the specialisation of the submitted tag, like: calculus, FERMAT, numerical, primes, etc. Note that the tag scientific programming has been detected as a specification of the tag math. This may at first appear strange but by considering that our folksonomy (de.licio.us) is populated mainly by computer scientists it could be expected. Finally, tags belonging to node 4 represent a sort of aggregation of the submitted tag, containing all the tags (terms) linked, or related to, the given tag, such as: visualisation, optimisation, 3D, and Homeschool.

In this paper, we present a different approach inspired by the semantic tagging on folksonomies that leads to a new methodology. This methodology is based on the semantic differential concept introduced by the psychologist and communication scholar (Osgood et al., 1975). The idea is to use the semantic differential concept approaching it by a completely different point of view from the original author’s proposal. The original work of Osgood considered the dimensions of meaning, using word semantics and highlighting the differences between individuals’ connotations for terms, and thus mapping the psychological difference between words. In his work, participants were given a word, for example ‘polite’ and prompted with a variety of adjectives to characterise it. These adjectives were displayed at both ends of a seven-point scale. The range of adjectives went from instance ‘good’ to ‘bad’, represented as opposite in a window and used to map individuals’ connotations for a given word. The opposite couples of adjectives used by Osgood are:

- angular/rounded
- weak/strong
- rough/smooth
- active/passive
- small/large
- cold/hot
- good/bad
- tense/relaxed
- wet/dry
- fresh/state.

We propose to use a clustering technique over the semantic differential approach for understanding the users’ perception of a website, or the perception of the institution or concept represented by that website. Following Osgood’s original work, we clustered the tags inserted by the users in three classes: evaluation, potency, and activity. We point out that in our work there are not fixed couples of opposite adjectives but we leave the users free of tagging a website with the adjectives they think best suits their opinions and evaluation of the interested website.

In fact, we believe that tagging systems can act as social dynamic enablers representing the real ‘vox populi’. Think about the metaphor of movie/theatres reviews: where reviewers (as domain experts) tend to highlight peculiarities, or evaluations, the average audience is not able to perceive or better perceive reviews in a completely
different way (not having the same cultural background). By employing a tagging system (folksonomy), we can leave the users free to express their own perception without restricting them in a frame like a questionnaire or a set of fixed couples of adjectives, thus capturing the real perception of the intended website audience. Moreover, by using collaborative tagging systems as a test bed for the users’ perception, we can detect tagging variations over time and update the perception classes according to the current users’ opinions on the subject matters.

4 Experimental results

To test our methodology of Semantic Differential through Collaborative Tagging, we built a basic collaborative tagging system.

The target of our experiment was to evaluate the perception of a community of users with respect to the web-portal of the Sapienza University of Rome (www.uniroma1.it). We decided to conduct this experiment as many users of the web portal complained about its features and usability. The community considered in the experiment is composed of three classes of users: students, academic and administrative staff. The study was conducted with 48 individuals taken among these three classes. The classes contain different ages and both males and females. Participants were invited to visit the portal and to tag it with a set of 3–5 adjectives. The system provided users with classic collaborative tagging functionalities, such as: presenting the document to be tagged and the text boxes to insert the corresponding tags. Figure 2 shows the tagging systems used for the experiment, although the system and the portal are in Italian, the tagging areas are clearly visible on the left-hand side.

**Figure 2** The collaborative tagging system used in the experiment (see online version for colours)
The majority of users were students about 60% from undergraduate to graduate levels. The remaining sample was split into 20% of administrative staff and 20% of academic staff subdivided into researchers and professors.

By using the collaborative tagging system, we collected around 162 tags from 48 individuals in a 2-weeks time frame. As a first step, we analysed the frequency of each tag (adjective). Table 1 summarises the most frequent tags, selected by computing a relative threshold (based on the mean value of the entire data set). We point out that there is a sort of antinomic behaviour among the general tags, looking at the positions in the table grouped by similar values, for example occurrence = 11, we found that two very different tags are the most frequently used. This behaviour works for almost every couple of tags in Table 1. Exploring the complete data set, we found this behaviour repeatedly, even with less frequent tags, and this lead us to suppose that even after clustering (assigning tags to the three classes: evaluation, potency and activity), the user perception will be split in mainly two neat categories according to the overall binary perception: positive or negative.

**Table 1** Table showing the most occurring tags

<table>
<thead>
<tr>
<th>Tag</th>
<th>Occurrence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simple</td>
<td>11</td>
</tr>
<tr>
<td>Dazed</td>
<td>11</td>
</tr>
<tr>
<td>Clear</td>
<td>6</td>
</tr>
<tr>
<td>Sad</td>
<td>6</td>
</tr>
<tr>
<td>Comprehensive</td>
<td>6</td>
</tr>
<tr>
<td>Inconsistent</td>
<td>5</td>
</tr>
<tr>
<td>Intuitive</td>
<td>4</td>
</tr>
<tr>
<td>Useless</td>
<td>4</td>
</tr>
<tr>
<td>Poor</td>
<td>4</td>
</tr>
</tbody>
</table>

We categorised the tags in the three classes according to the clustering proposed by Osgood in Zipf (1949), as shown in Table 2. The categorisation was carried out by associating to each tag the class by using a combination of thesauruses and dictionaries. The same approach has been employed for the positive or negative meanings of a tag. The criteria to assign a tag to a class are:

- **Evaluation**: representing the overall feelings about the website (adjectives like good or bad)
- **Potency**: representing the expressive power and impact on the perception of the website (adjectives like strong or weak)
- **Activity**: representing the possibilities and functionalities (at informational level) offered by the website (adjectives like: active or passive).
By using the clustering results, we analysed the perception of each user, expressed by the tags provided to the system with respect to the selected classes (evaluation, potency and activity) and their positive or negative implications.

Figure 3 shows the sample identified by progressive integer numbers 1–47.\textsuperscript{1} The three colours represent the classes, while each sector displays the normalised score and category to which each tag inserted by the users belongs to (radar view).\textsuperscript{2} From the Radar graph, we can extrapolate that the activity class has a trend towards negative values and, in general, the graph is distributed among negative values. Furthermore, this graph highlights that users with a positive evaluation of the website focus their attention to that particular class, which deals mainly with strong feelings about a website (adjectives like: good or bad, nice or ugly, etc.). Finally, Figure 4 shows a three-dimensional view of the distribution of the entire data set. Each point in the scattered 3D plot represents the values (positive or negative) normalised between $[-1, 1]$ of the tags among the three classes represented on the three axes. Note that, there is quite a big cloud of points where all the values are negative or where only evaluation and activity are positive. This behaviour represents the perception of the website among the community of users, which is generally negative and positive only for users ignoring the potency (expressive power and impact) of the website.

Table 2 Classes assigned after the clustering phases and number of tags falling in their relative class; the third rows represent the number of tags per class with positive and negative meanings, (+$n$) stands for $n$ tags in that category with positive meaning and ($-n$) for the negatives.

<table>
<thead>
<tr>
<th>Evaluation</th>
<th>Potency</th>
<th>Activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>82(+40)(–42)</td>
<td>52(+11)(–41)</td>
<td>28(+8)(–10)</td>
</tr>
</tbody>
</table>

Figure 3 A radar view of the users with respect to the selected classes (evaluation, potency and activity) (see online version for colours)
5 Conclusions and future work

In this work, we explored a methodology for detecting the users’ perception of web services by using collaborative tagging systems. The approach highlighted the potential of such systems in the evaluation of the end users perception without requiring questionnaires or very structured approach like the semantic differential technique.

To test our approach, we built a basic collaborative tagging system used for the evaluation of the users perception in the context of a university web portal (www.uniroma1.it). We collected data from 48 end users (students, academic and administrative staff) in a short-time frame. Our hypothesis, after conducting some initial interviews, was of an overall negative perception of the university website. The results obtained from the collaborative semantic approach confirmed our hypothesis, and gave us additional information. The overall perception of the users seemed to be split into mainly two different points of view: on the one hand, many users had a deep negative impression of the website, along all the three classes (evaluation, potency and activity) used for the evaluation; on the other hand, another set of users disclosed a positive attitude towards the website considering only the activity and estimation classes.

These experimental results are encouraging in the sense of exploring collaborative tagging systems as an alternative in designing a system for evaluating the users’ perception of websites (even if it demonstrates to be general enough for being applied to different domains).

We think that the overall methodology can be further automated by employing clustering techniques such as $k$-means applied over a broad collaborative tagging system like del.icio.us. We intend to experiment in future works this automation process. One approach can be of exploring the del.icio.us folksonomy to capture the users’ perception of a tagged website, and to use co-occurrences information (by using a SemanticHalo like approach for resolving synonymy and polysemy) as a distance measure among the users’ tags and basic sets of opposite adjectives falling in the three classes isolated in this study. These distances can then be used for computing a $k$-means clustering, so as to also automate this phase.
References


Notes

1 One user has been deleted from the sample because inserted tags as spam, due to the anonymous login to the system.

2 A radar graph, sometimes called a star or spider graph, is laid out in a circular fashion, rather than the more common linear arrangement. A radar graph consists of axis lines that start in the centre of a circle and extend to its periphery. Each axis can either represent an independent measure related to a single thing or a single measure broken into multiple subdivisions of a single category.