Tags, Networks, Narrative: Exploring the use of social software for the study of narrative in digital contexts

Bruce Lionel Mason De Montfort University The Gateway Leicester, LE1 9BH, UK +44 (0)116 250 6509 bmason01@dmu.ac.uk Sue Thomas
De Montfort University
The Gateway
Leicester, LE1 9BH, UK
+44 (0)116 250 8266
sue.thomas@dmu.ac.uk

"Tags, Networks, Narratives" is a project that assessed the potential for innovative, online social-software in connecting academics across disciplines. As part of the research, 33 volunteers spread across the world were asked to use Del.icio.us to tag 40 sites. Once all sites were tagged, the volunteers were asked to reflect on the process.

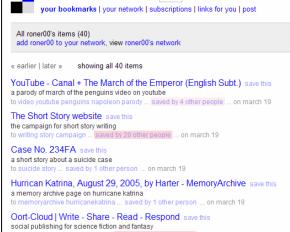
Tagging study procedure

- 1. 33 participants; 28 completed all or most of the study.
- 2. Tag 40 sites in batches of 10.
- 3. Break participants into 5 groups.
- 4. Inspect other group members' tags.
- 5. Derive up to 5 "key" tags for group.
- 6. Re-tag 5 of the sites as a group.
- 7. Answer follow up questions by email.

Tagging Study aims

• To create a "tagverse" of at least 1000 tagging instances for a fixed number of sites.

- To compare the resultant folksonomy with a standard taxonomy.
- To investigate taggers' choices of tags.
- To investigate taggers' perception of others' tags.
- To investigate whether participation in the study changed how the taggers' perceived the act of tagging.
- To investigate implications for "best practice" in tagging.



del.icio.us / roner00 /









Tags, Networks, Narrative: The tagverse

- 1396 unique tags used by the taggers.
- 1233 "cleaned" tags.
- 1062 "processed" tag.

A cleaned tag is one that is corrected for obvious misspellings and punctuation problems and where capitalization is made consistent.

A processed tag is one which judgment calls about ambiguity and intent are made.

This folksonomy was then processed through Wordnet which found 717 unique words that matched at least one tag that had been used by the taggers.

According to Jennifer Trant's preliminary analysis of data from the Project Steve social tagging study, social tagging generated approximately 70% more terms than the taxonomy which was already in place. In this study, the folksonomy generated:

- 1396 unique tags: 95% more tags than Wordnet.
- 1233 "cleaned" tags: 72% more tags than Wordnet.
- 1062 "processed" tags: 48% more tags than Wordnet.

Some tagging issues

- Is "blog" different from "blogs"?
- What is the difference between ww2 WW2 wwii worldwar2 World-War-2 and so on?
- Is a shortstory a story?

Tagging patterns

- Narrow tagging vs. broad tagging
 - One person used 267 tags (for just 40 sites).
 - Most had around 100-120

It appears that there were two basic types of tagger. One was focused on indexing the sites and, therefore, attempted to create a consistent, small, number of tags. The other type used idiosyncratic, descriptive tags.

Want to know more?

- End of project seminar, De Montfort University, Leicester, UK, 11am, September 18th, 2007.
- Free public lecture about folksonomy by Thomas Vander Wal, De Montfort University, Leicester, UK, 2pm, September 18th, 2007.

Visualizing the tagverse







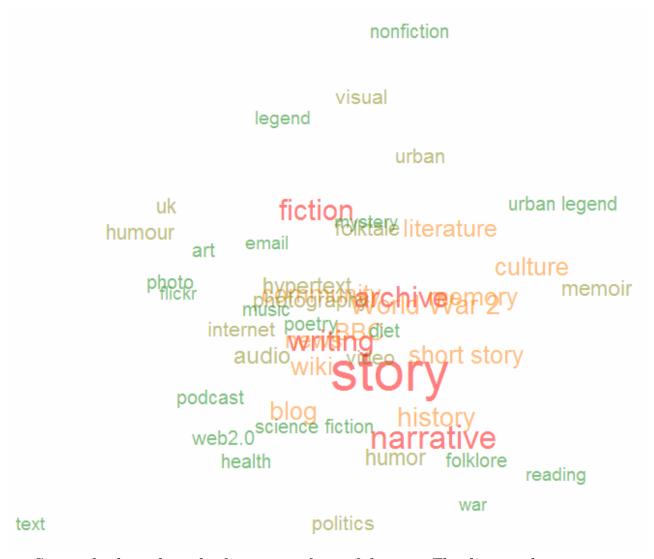
Note some of the choices made by three of the taggers.

- Boyil74 uses compound words, cabes39 uses hyphens, cabir98 uses neither. Thus boyil has the tag martinlutherking while cabir98 has three tags: martin, luther and king. (Cabes39 did not use this tag when he tagged the page it is relevant to).
- Cabes 39 has idiosyncratic tags (bumpf) and idiosyncratic usage: fossil is used to refer to a website that is no longer being updated.
- The most popular tags for the three users are significantly different.
- Boyil74 uses story where cabir98 uses stories. Cabir98 uses penguin where boyil74 uses penguins.

A tagger speaks

I enjoyed more the taggings that are accompanied by a note. Otherwise, I feel that the tag itself doesn't give me enough information to become interested in their web sites. Also the tags can be very personal, which is great. For example someone has a tag called "Calvino", but the web site is not about Italo Calvino. The web site makes her think of Calvino, it shows the mental links she has made, which I consider great, and it is an interpretation of the web site. I think this is one of the values of tagging: interpretation and mental links, rather than categorisation because of the sake of categorisation, does it make sense?. Also, I liked when people tag with composed words such as "truestories", "creationmyth", because there is a bit more of information to become interested. Also I like to find a short collection of tags. I feel lost when there are too many.

Dynamically visualizing the tagverse



Size and colour show the frequency of use of the tags. The distance between tags shows how related they are. The relatedness is a function of how many shared users the tags have.

Notes

If folksonomy is to be used in academic communication across disciplines then careful attention will need to be paid to the productivity of tagging and the diverse styles, motivations and levels of expertise of individual taggers.

^[1] Trant, J. Social Classification and Folksonomy in Art Museums: early data from the steve.museum tagger prototype. *Proceedings of the ASIST-CR Social Classification Workshop*, November 4, 2006.