Tagging YouTube - A Classification of Tagging Practice on YouTube

Stacey Greenaway[1] , Mike Thelwall[2] and Ying Ding[3]

[1] *s.greenaway@wlv.ac.uk*
School of Computing and Information Technology, University of Wolverhampton, Wolverhampton (UK)

[2] *M.Thelwall@wlv.ac.uk*
School of Computing and Information Technology, University of Wolverhampton, Wolverhampton (UK)

[3] *dingying@indiana.edu*
School of Library and Information Science, Indiana University, Indiana (USA)

Abstract
A problem exists of how to categorise the abundance of user generated content being uploaded to social sites. One method of categorisation being applied is tagging, user generated keywords that are assigned to the content. This research presents a study into the tagging practice of YouTube users. A classification scheme was applied to a dataset of 768 tags, assigning the tags to different categories of tag type. Analysis reveals how useful the tagging method on YouTube is at improving the categorisation of user generated video content in contrast to collaborative tagging systems.

Introduction
One of the key attributes of Web 2.0 sites, in conjunction with social networks, is tagging; a useful tool for labelling online resources like web pages, audio, images and video. This article is primarily concerned with online video. Video sharing websites such as YouTube, GoogleVideo and MySpace.tv provide User Generated Content (UGC) to mass audiences. The diversity of user generated video creates difficulties for categorisation and findability. Current methods of searching for internet video using existing keyword search techniques are inadequate because of the lack of meta data available for videos. Titles, descriptions, social information and minimal existing tag data are insufficient to accurately describe the content of video (Yang et al. 2007).

At present tagging on YouTube is not collaborative, with only the owner of the video being able to tag. If collaborative tagging was introduced, any user could tag any video, increasing the number of tags for a video and potentially improving video search. Geisler and Burns (2007) found that YouTube tags added additional description of the video content that was not found in other text on the page. Halvey and Keane (2007) found that few users interact with the social element of the site i.e. join groups, upload videos, make friends, favourite videos or comment. If few users upload content to many viewers, only the tags of a few users are being used as additional textual data for the videos. Therefore, whilst there is potential for more tags to be entered and for rich folksonomies to be created, there remains the problem that if only a few users interact with the social elements of YouTube, how can users be encouraged to tag?

This research presents an analysis of tagging behaviour on YouTube, through a classification of the user generated tags assigned to a random selection of 100 YouTube videos. Tags were classified into various categories of tag type, using a custom classification scheme. The research investigates whether theories of structure, motivation and tag type applied to collaborative tagging systems (Golder & Huberman, 2006, Marlow et al. 2006, Angus et al. 2008) are evident in YouTube tag data. It is a preliminary study into understanding how useful the tags entered by the uploader of the video are at describing the content to other YouTube users and if the absence of collaborative tagging has an impact on the types of tag and the cognitive level of the vocabulary.

Proceedings of ISSI 2009, pages 660–664.
Edited by B. Larsen and J. Leta.

Background

Golder and Huberman (2006) claim that the main problem with tagging stems from its free-form nature. The absence of any controlled vocabulary means that tags have a multitude of different spellings, plurals, terminology, opinions, descriptions, dialects and languages. Croft and Cruse (2004) argue that words can be categorised based on their level of specificity, or cognitive level. When applied to tags, there are three cognitive levels superordinate, basic and subordinate. Basic level tags have the least cognitive cost to the user – that is they are thought of more quickly. They are more likely to have a high frequency as there is a greater probability for agreement of terms than subordinate level tags (Golder & Huberman, 2006).

Collaborative tagging of images on the Flickr website has provoked research into tagging behaviour, types of tag and semantic relations, (Aurnhammer et al., 2006; Marlow et al., 2006; Rafferty & Hidderly, 2007; Ames & Naaman, 2007, Angus et al., 2008). The research has revealed the quantity and diversity of tags entered by both resource owner and other users. Research into tagging on YouTube is not as extensive as that of Flickr. Research centres on quantitative analysis of YouTube tags (Geisler & Burns, 2007; Halvey and Keane, 2007; Paolillo, 2008) rather than focussing on the vocabulary of the tags.

Ding et al. (submitted) highlight a problem with analysing tags in YouTube; because only the user uploading the video can tag, there is no indication of the collaborative opinion of viewers of the video. YouTube tags can only indicate trends in the type of content being uploaded to the site, but can not offer insight into the type of content users prefer watching. The authors note that using tag frequency to identify community interest is not possible in YouTube.

Methods

Data Collection

The dataset of Ding et al. (2008) was used for this study. The data was originally collected as follows: In September 2007 a crawl of YouTube was conducted to obtain a dataset of video URLs and tagging data. The crawler started from the main page at http://youtube.com and visited every available video page (links starting with http://www.youtube.com/watch?v). On each video page it collected tagging data and visited the links pointing to other video pages. YouTube does not provide related tag data. In order to avoid visiting the same page more than once, the query parts of links were ignored.

The original dataset contained 43,641 tags. The majority of foreign words or characters in particular, Chinese/Japanese characters that had not converted correctly into the text file were manually removed; 1,461 entries were removed leaving a dataset of 42,180 tags. A random selection of 100 videos and their assigned tags were then extracted from the dataset using a custom script. This created a dataset of 768 tags for Classification.

Classification Scheme

Angus et al. (2008) developed a classification scheme based on possible image categories in Flickr. For the purposes of this research, the classification scheme was modified to be more suited to a classification of YouTube Tags. The distinction between social and personal motivation was removed, with categories in A and B being tags generally descriptive of the content and categories in C being of use only to specific users or groups within the YouTube community. Rather than miscellaneous categories as defined by Angus et al. (2008), categories in D are tags which are either irrelevant, or seen as not useful in terms of describing or indentifying the video in search or tag browsing. Alongside restructuring the classification scheme, five new categories were added.

Tags were classified whilst watching the respective video, by a single classifier. Some videos were no longer available, and so the tags assigned to these videos were classified into the D5 (unable to determine relationship) category.

Findings

A large number of tags referred to people, this is not depicted by the B1b (people/animals/objects) result of 9.5% as the majority of these tags were classified into the D2 (Multi Words) category. The largest percentage of tags, 23.3%, were placed into the D2 category. Some of the tags classified in this category resulted from complete sentences being placed in the tag field, either as a description of the content or the title. Considering this tagging practice by users, a surprisingly low result of 3.3% was recorded for the D7 (Conjunctions and Prepositions) category. It had been expected that a higher percentage of these tags would be found in relation to the other categories, due to the finding in Ding et al. (submitted) that '*the*' is the most frequently assigned tag for the years 2006 and 2007 and fourth in 2005.

Table 1 – Total number of tags and corresponding percentage of all tags, for each classification category.

	Classification Category	No of tags	%age of all tags
A	**Generic relationship between tag and video content**		
A1	Tag generically identifies what video is 'of'	85	11.1%
A2	Tag identifies video Category/Genre	42	5.5%
B	**Specific relationship between tag and video content**		
B1a	Tag specifically identifies what video is 'of' (place names/events)	66	8.6%
B1b	Tag specifically identifies what video is 'of' (people/animals/objects)	79	9.5%
B2	Tag identifies what video is 'about'	67	8.7%
B3	Tag identifies opinion expression	51	6.6%
C	**Tag only useful to a minority of users, specific individual or group**		
C1	Refining tag	45	5.9%
C2	Self-reference tag	5	0.7%
C3	Tag which explicitly denotes ownership of video	8	1%
D	**Irrelevant/Non Useful Tags**		
D1	Compound tag (truncating or compounding words to form one tag)	3	0.4%
D2	Multi-word tags (individual words in these)	179	23.3%
D3	Attention attracting tags	3	0.4%
D4	Misspelling	4	0.5%
D5	Unable to determine relationship	39	5.1%
D6	Foreign word/character	67	8.7%
D7	Conjunctions and prepositions	25	3.3%

Category A1 (what the video is of) and A2 (category/genre) will contain mostly basic level tags that describe the content at its most general. 11.1% of all tags were classified A1 and was the second highest category. Surprisingly, A2 contained only 5.5% of tags, suggesting that YouTube taggers describe the video content more than they use tagging to categorise the video, using the pre-assigned YouTube categories only. This finding is emphasised by the high percentage of Category B tags, that more specifically describe the video content and may require some specialist knowledge. B1b (9.5%), B2 (what the video is about) contained 8.7% of tags, B1a (places/events) contained 8.6% and B3 (opinion expression) 6.6% of all tags. An indication that YouTube taggers use more specific level vocabulary over basic level generalised terms is that 5.9% of tags were classified as C1 (refining tag) tags. The tendency of YouTube taggers to use more subordinate level, descriptive tags could explain the low percentage, 0.4% of category D3, attention attracting tags. It would be expected that these tags would be of basic level vocabulary, maximising the probability of agreement on terms, with tags being words that are perceived to be regularly searched for, or relate to popular categories or videos. To accurately assess the specificity of the tag vocabulary, tag frequency and co-occurrence metrics can be analysed (Golder & Huberman, 2005; Cattutto, 2007). This is not possible with this data sample as only 6.6% of tags occur more than once.

Discussion

Collaborative tagging allows for the taggers in the system to classify and categorise the content in the system using language useful to the community. In YouTube this doesn't exist, as only the owner of the video can tag and they may not use language or a style of tagging that is useful to the community. Without collaborative tagging there is no agreement between taggers that tags are good, useful and relevant to the content and as a result there is no reuse of tags by which to measure tag relavance. More multi word tags were identified than compound tags, this can be seen as a positive tagging behaviour of YouTube users. Multi word tags are more useful in keyword search than compound tags as users are unlikely to enter the compounded word as a search term. Multi word tags can also be useful to create long-description meta data for videos that can improve indexing of videos. However there is a usability problem of how to accept and handle multi word tags in a tagging system.

Conclusion

The results suggest that YouTube users use tagging as an extension of the description and title fields. Tags do not appear to be used to further categorise a video, with users apparently relying on the categorisation structure of the YouTube system for this purpose. This is surprising since Flickr tags seem to be frequently useful for this purpose (e.g., Angus et al., 2008) and suggests that YouTube video posters are less aware of the need to publicise their work through tags. The classification found that YouTube taggers used a relatively specific vocabulary to describe their videos, for instance, tagging the species of dinosaur, rather than just tagging dinosaur; or tagging the make and model of motorbike, as apposed to just entering the motorbike tag. Whilst these tags may be useful at finding less popular videos through keyword search, in theory, searchers are unlikely to use more specific vocabulary for keyword terms, so the tags may well be relevant to only a few users rather than the majority (Furnas et al., 1987; Golder & Huberman, 2006). It may not be the case that the syntax used is too specific for the majority of users, but rather that without the collective vocabulary provided by collaborative tagging it is impossible to accurately assess the specificity of the tags or the level of agreement of terms achievable. The lack of agreement between YouTube tags makes the clustering of videos for related content impossible, impacting on their potential for categorising user generated videos.

Through analysis and classification of collaborative tagging data it is possible to evaluate the collective intelligence of the community, to assess the social impact of a resource or user, to discover community interest, trends, popularity and social connections. The method of tagging implemented in YouTube does not allow for such evaluations, and it is not clear why this is the case. With the introduction of a collaborative tagging system it would be possible to assess the popularity of the videos through analysis of the amount of tags entered per video, the type of tag entered, language used and opinions expressed. Trends in viewing habits could be uncovered, which could improve the recommendation of videos. Recommendation systems could be developed based on shared user interest and co-occurrence of tags. The tags themselves could provide a method for categorising the increasing amount of user generated content, either for retrieval, for curating collections, or for preservation of content.

References

Ames, M. and Naaman, M. (2007). Why we tag: motivations for annotation in mobile and online media. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (San Jose, California, USA, April 28 - May 03, 2007)

Angus, E., Thelwall, M. & Stuart, D. (2008) General patterns of tag usage among university groups in Flickr. Online Information review, 32(1), 89-101

Aurnhammer, M., Hanappe, P. & Steels, L. (2006) Integrating collaborative tagging and emergent semantics for image retrieval. Proc. of the Collaborative Web Tagging Workshop (WWW '06).

Cattuto, C., Loreto, V. & Pietronero, L. (2007) Semiotic dynamics and collaborative tagging. *Proceedings of the National Academy of Sciences (PNAS),* **104**(5)**,** (pp. 1461-1464).

Croft, William A. & D.A. Cruse (2004). Cognitive Linguistics. Cambridge: Cambridge University Press.

Ding, Y. Toma, I. Kang, S. Zhang, Z. and Fried, M. (2008). Mediating and Analyzing Social Data. Proceedings of The 7th International Conference on Ontologies, Databases, and Applications of Semantics (ODBASE 2008), Lecture Notes in Computer Sciences, Nov 11-13, 2008, Monterrey, Mexico, Springer-Verlag.

Ding, Y., Jacob, E.K., Zhang, Z., Foo, S., George, N.L., Guo, L. & Yan, E. (submitted) Adding Semantics to Social Tags.

Furnas, G.W., Landauer, T.K., Gomez, L.M., & Dumais, S. T. (1987). The vocabulary problem in human-system communication. Communications of the ACM, 30(11), (pp. 964-971).

Geisler, G. & Burns, S. (2007) Tagging video: conventions and strategies of the YouTube community. JCDL '07: Proceedings of the 2007 conference on Digital libraries, (pp. 480-480)

Golder, S. & Huberman, B. (2006) Usage patterns of collaborative tagging systems. Journal of Information Science, 32(2):198--208.

Halvey, M. J. & Keane, M. T. (2007). Analysis of online video search and sharing. In Proceedings of the Eighteenth Conference on Hypertext and Hypermedia (Manchester, UK, September 10 - 12, 2007). HT '07. ACM, New York, NY, 217-226

Marlow, C., Naaman, M., Boyd, D. & Davis, M. (2006) HT06, tagging paper, taxonomy, Flickr, academic article, to read. HYPERTEXT '06: Proceedings of the seventeenth conference on Hypertext and hypermedia, (pp. 31-40).

Paolillo, J. C. (2008). Structure and Network in the YouTube Core. In Proceedings of the Proceedings of the 41st Annual Hawaii international Conference on System Sciences (January 07 - 10, 2008). HICSS. IEEE Computer Society, Washington, DC, 156.

Rafferty, P. & Hidderly, R. (2007) Flickr and Democratic Indexing: Dialogic Approaches to Indexing. Aslib Proceedings **59**(4) (pp.397-410).

Yang, L., Liu, J., Yang, X., & Hua, X. 2007. Multi-modality web video categorization. In Proceedings of the international Workshop on Workshop on Multimedia information Retrieval (Augsburg, Bavaria, Germany, September 24 - 29, 2007).