

Tagging YouTube - A Classification of Tagging Practice on YouTube

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

Tagging as a method of assigning keywords to digital resources has been used as a personal organisation tool for the desktop for a number of years (Vander Wal 2007). It is the adoption of the process on the Internet and the transformation from personal organisation to social application that has introduced tagging to a wide audience and increased its popularity. Seen as one of the key attributes of Web 2.0 sites, in conjunction with social networks, tagging is a useful tool for labelling online resources like web pages, audio, images and video.

The rise in popularity of social or collaborative tagging systems such as del.icio.us and Flickr has led to analysis of the tags created by these systems. A term coined by accident through discussion of tags as a folk taxonomy: folksonomy means an informal taxonomy created by folk, and is often credited to Vander Wal (2007) although Mathes (2004) debate whether the term should be used. Despite such debate, folksonomy has been adopted as the standard term to describe tag data and how it categorises the tagged resources. The categorising of resources is determined by users' opinions of the resources, as opposed to system developers or experts. Folksonomies are not hierarchical and the free association of tags provides an abundance of additional content, or meta data for an individual resource. Tagging and Folksonomy research is largely concerned with semantic similarity and lexical relations. Folksonomies provide useful social information about the various ways different users tag the same resource, they offer information on dialect, vocabulary, slang and how tags are affected by the social network.

This article is primarily concerned with online video. Collaborative tagging can be useful for improving indexing and categorising internet video by providing a user generated text alternative to the visual content. As a low cost method of video annotation, collaborative tagging has value because it is created by viewers of the video. The tags are potentially accurate and descriptive of the content as well as expressing opinions about the video. Collaborative tagging can provide a social commentary about a video (Shamma et al. 2007). With the potential for international audiences, the resulting descriptions can make the video more widely accessible. Tags can be unbiased; therefore unlike keywords generated by the creator of the video, they can express a wider range of opinion and offer a reflection of the content of the video from a wider range of perspectives. The complexity of visual images can still be best captured and recorded by human interpretation, and tagging offers a method for each viewer of the video to annotate the video with their interpretation of the content.

Current methods of searching for internet video using existing keyword search techniques are inadequate because of the lack of meta data available for videos. Yang et al. (2007) highlight that titles, descriptions, social information and minimal existing tag data are insufficient to accurately describe the content of video. Browne and Smeaton (2004) identify problems in video retrieval; automated indexing is only able to identify low-level features such as colour or the human face, but high-level features such as recognising the person in the video requires

human recognition. Collaborative tagging is potentially useful as a partial solution to this problem as users can identify and tag high level features of videos. Rafferty and Hiderly (2007) discuss high level and low level features and how collaborative tagging can bridge the gap between the automatic extraction of low level features to the high level descriptions provided by human content. Halvey and Keane (2007), in a study of YouTube, found that more descriptive information about a video correlated with more views. This is explained by the fact that search engines use text matching techniques to find videos; therefore, the more textual information surrounding the video, the higher the probability of it being returned by a query. This supports the idea that increasing the number of tags for a video will improve video search. The best way to increase the number of tags is to introduce collaborative tagging.

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 (Yang et al. 2007). Cha et al. (2007) highlight the problem of finding niche videos in UGC video on demand sites like YouTube, when popularity is ranked by number of views, and popularity follows a power law, with the same few popular videos being watched by the majority of users; a trend also observed by Gill et al. (2007). Paolillo (2008) found, through analysis of social groups in YouTube, that users in groups did not link to the most popular videos. Whilst videos uploaded or 'favourited' by social groups could then be seen as key to recommending niche videos to more users, 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. At present tagging on YouTube is not collaborative, with only the owner of the video being able to tag. 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. If collaborative tagging was introduced, any user could tag any video, 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. This is designed to help understand which attributes comprise a useful tag. The work is a preliminary study to gain a more detailed understanding of the tagging behaviour of YouTube users and the types of tag they enter, extending the findings of Ding et al. (submitted). In order to facilitate further research into a suitable method of collaborative tagging, that encourages users to tag YouTube videos with tags that are useful at improving categorisation and search.

Background

The structure of collaborative tagging systems has been analysed by Marlow et al. (2006), Golder and Huberman (2005), Lambiotte (2005) and Voss (2007). These view tagging systems as a tripartite network with the relationship between user, resource and tag being central. Marlow et al. (2006) provide a taxonomy from which all tagging systems can be assessed. Voss (2007) extends the taxonomy adding further attributes, Sen et al. (2006) although not referencing Marlow et al (2006), discuss attributes of tagging systems in relation to their Movielens system and further extend the taxonomy for the design of a tagging system.

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. This vocabulary problem was first identified by Furnas et al. (1987). It relates to semantic differences between word usage and ultimately how this effects keywords (or tags) application in information retrieval. Golder and Huberman (2005) develop the ideas presented by Furnas et al. (1987) and describe the semantic problems generic to tags: synonymy, polysemy and homonymy. They also distinguish between cognitive levels of tags describing how basic level tags have a greater probability for agreement of terms than subordinate level tags. Much subsequent research has gone on to develop the ideas presented by Furnas et al. (1987) and describe the semantic (Enser & Sandom, 2007; Aurnhammer et al., 2006) and cognitive (Golder & Huberman, 2005; Sinha, 2005; Trant, 2006; Cattuto, 2007) problems which are generic to tags.

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 more likely to be agreement on basic level tags. Superordinate and subordinate have a higher cognitive cost. In relation to collaborative tagging – superordinate level is difficult to assess Golder and Huberman (2005) and Cattuto (2007) differentiate only between basic and subordinate level tags.

The tag types defined by Golder and Huberman (2005) are: *what or who it is about, the type of resource, who owns it, qualities or characteristics, refining categories, self reference and task organizing*. These types can be categorised as personal or social. Marlow et al. (2006) suggest that these tag types are generated out of organisational (which can produce both personal and social tags) and social motivations. A personal organisational motivation would be *future retrieval*, the social variant being *contribution and sharing*. Social motivations include: *attract attention, play and competition, self presentation and opinion expression*. Differing motivations will affect the types of tags entered. Zollers (2007) extends the research of Marlow et al. (2006) and Golder and Huberman (2005) by analysing the tag data from a selection of collaborative tagging systems for evidence of the tagging motivation and the type of tag. Motivations to tag video resources in del.icio.us are discussed by Paolillo and Penumarthy (2007) with reference to user incentives defined by Marlow et al. (2006). The main motivation when using del.icio.us is one of future retrieval, this may result in a set of more personal tags as opposed to socially motivated tags. Ames and Naaman (2007) extends the work of Marlow et al. (2006) through user interviews to provide evidence for the tagging motivations they suggest. Predominance for social over personal motivation was found, the most popular motivation being to enter tags to improve findability and describe content. However, whilst being a social motivation, it is fuelled when the resource holds personal relevance to the tagger. Thom-Santelli et al. (2008) researched social tagging motivations and categorise the tagging roles of users of social tagging systems, as opposed to the tags themselves. They suggest that an understanding of tagging motivation and behaviour could lead to tag recommendation systems, suggesting tags used by users with similar motivations. Similarly, Santos-Neto et al. (2007) see better understanding of the tagging behaviour of users to recommend content rather than tags. Li et al. (2008) propose a method of recommending content and potential friends to users by measuring co-occurrence of user generated tags and clustering users based on common interests, rather than any social connections they may already have.

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 & Hilderly, 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. Geisler and Burns (2007) published findings of a quantitative analysis of YouTube tags. They found a mean number of tags per video to be 6 and that 66% of the tags added additional description of the video content that was not found in other text on the page, such as title, description or author.

Halvey and Keane (2007) conducted a study of YouTube focussing on search and user behaviour. They found that most users only use YouTube to search and watch videos, but few users interact with the social element of the site i.e. join groups, upload videos, make friends, favourite videos or comment. They state a potential for projecting the data of *Active users* onto *Passive users*. If collaborative tagging was implemented on YouTube, passive users would benefit from the tags entered by the active users. A further study found if very few tags were entered then the video received very few views. The average number of tags per video, for their dataset, was found to be 4.1 with the maximum amount entered being 25. However, for videos that are recommended by YouTube on the front page, the average number of tags was found to be double that at 8.73%. The authors found no evidence that considerably increasing the amount of tags beyond the recommended video average substantially increased the views.

Both Capra et al. (2008) and Paolillo (2008) harvested a data set of YouTube videos and the related textual data by crawling YouTube through links in blog posts. Whilst the aim of Paolillo (2008) was to find popular videos in the social core of YouTube, Capra et al. (2008) chose a method to find videos related to the US Election 2008, with their focus being how to curate user generated video collections: how to find, filter and ultimately preserve important, relevant and interesting videos. Paolillo (2008) characterised users based on the user generated tags and user profiles collected in the crawl and clustered users based on their usage of high frequency (commonly occurring) tags. The study found users in social groups on YouTube were more likely to provide links to less popular videos, as no high ranking/most popular videos were found in the dataset.

Ding et al. (submitted) analysed tags and tagging behaviour as part of a comparison study of social tagging over three social networks, del.icio.us, Flickr and YouTube. The popularity of tags over time was analysed by comparing the most popular tags and tagging behaviour over two years, 2005 and 2007. By comparing tag popularity over time emerging trends in topics of interest were revealed. The study highlights a problem with analysing tags in YouTube, as 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.

Research questions

YouTube is a social site that allows users to form friendships, share video and interact with the videos of other users through comments and video responses. Unlike other social sites, such as Flickr and del.icio.us, YouTube does not allow users to tag other users' videos. Tagging is therefore not collaborative, but performed by the video owner. This research investigates whether theories of structure, motivation and tag type applied to collaborative

tagging systems (Golder & Huberman, 2005, Marlow et al. 2006, Angus et al. 2008) are evident in YouTube tag data. As a preliminary study into understanding which attributes construct a useful set of tags in terms of improving the categorisation of user generated video, the following two research questions are addressed.

1. How useful are the tags entered by the uploader of the video at describing the content to other YouTube users?
2. Does the absence of collaborative tagging impact on the types of tag and the cognitive level of the tag vocabulary?

Methods

Data Collection

The dataset of Ding et al. (submitted) 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 (i.e. <http://www.youtube.com/watch?v=X2IExa2A198> and http://www.youtube.com/watch?v=X2IExa2A198&watch_response lead to the same video).

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, notions of “of” and “about” (Shatford, 1986, 1984 in Angus et al., 2008) and the notions of tag type defined in Golder and Huberman (2006). Categories in the scheme were further grouped based on social or personal motivation to tag. For the purposes of this research, the classification scheme was modified to be more suited to a classification of YouTube Tags. As tagging on YouTube is primarily socially motivated and carried out by the uploader of the video, there was no requirement for the task organising category (e.g., tags such as toread, toprint, towatch). Angus et al. (2008) found no task organising tags in the Flickr study, an assumption was made this would also be true 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 identifying the video in search or tag browsing.

Alongside restructuring the classification scheme, five new categories were added. With the addition of category A2, a distinction was made between tags that identified generically what the video is of and that identified a YouTube category or asserted a genre e.g. Comedy,

Music, Horror, Rock. Category B2 was created for tags that expressed an opinion about the video as a whole or certain qualities and characteristics of the video such as funny and scary. Three categories were added to account for irrelevant tags. D2, (multi-word tags) handles tags that only have meaning when viewed in context with the other tags assigned to the video, often names, titles and descriptions are entered as tags, but as single word tags viewed in isolation the tag becomes meaningless. This practice leads to an abundance of conjunctions and prepositions (e.g., the, in, of, and) a separate category D7 was created to handle these tags. Category D3, attention attracting tags, was added from an assumption that some users uploading videos have a primary motivation to tag to get more views for their video and would therefore add tags containing popular search terms (e.g., porn, sex, celebrity name) in order to achieve this. Table 1 below shows the classification scheme used and explains each category.

It was important to classify the tags whilst watching the respective video in order to correctly ascertain the meaning of some of the tags. For instance, it is difficult to classify C3 (denotes ownership) tags without first visiting the video page to find the username of the uploader. This fact questions the usefulness of some tags, as to be useful for search and discovery of video, they need to be meaningful in isolation from the content. Some videos were no longer available, and so the tags assigned to these videos were classified into the D5 (unable to determine relationship) category. Despite it being possible to classify some of the tags, a decision was made that the tags could not be accurately classified without watching the video first.

Table 1. Classification Scheme, Category Definitions - adapted from (Angus et al. 2008)

| A | | Generic relationship between tag and video content |
|----------|------|---|
| | 1 | Tag identifies what video is of at its most primary and objective level - no subject specific knowledge is needed to make this distinction (e.g., a video of a cat, tagged as 'cat' or 'animal'). Also included is the tag video. |
| | 2 | General YouTube defined Category or Genre (e.g. Comedy, Entertainment, Music) |
| B | | Specific relationship between tag and video content |
| | 1(a) | Tag identifies what video is of . Familiarity or some existing knowledge is needed to make this connection, and to a certain extent an assumption has to be made about this connection. Tags which identify place names/events – a video of a concert tagged with the band name and venue, or a football match tagged with the team name, or an individuals holiday video tagged with the destination, requires knowledge acquired from familiarity with the specific place/event in question. Assumptions have to be made that a video tag is what it claims to be if the video is not familiar. |
| | 1(b) | Tags which identify people/animals/objects – a video of Elvis Presley tagged as 'Elvis Presley' requires knowledge and familiarity of Elvis Presley. Distinctions cannot always be made between 'famous' people and 'non-famous' people, therefore the assumption has to be made that a video of a girl, tagged as 'Sarah' is in-fact a video of a girl who is called 'Sarah'. |

| | | |
|----------|---|--|
| | 2 | Tag identifies what the video is about Typically expressed by the use of abstract nouns or adjectives - an interpretation is made of what the video is about (e.g., video of people smiling tagged as 'happiness'; video of cars on a motorway tagged as 'speed'). |
| | 3 | Tags which express opinion of the content Includes Golder and Huberman (2005) tag types of <i>Qualities and Characteristics</i> and <i>Opinion Expression</i> (e.g. 'funny', 'rubbish') |
| C | | Tag only useful to a minority of users, specific individual or group |
| | 1 | Refining tag Tag which cannot stand alone - only useful when looked at as part of the larger tag set (e.g., episodes of a series of videos specified by a number; acronyms or dates.). |
| | 2 | Self-reference tag Tags which identify video content in terms of it's relation to either the tagger or the specific group which the video belongs to (e.g., 'my dog'; 'our graduation') OR tags which appear useful, but show no relationship/connection to the accompanying video. |
| | 3 | Tag which explicitly denotes ownership of video (e.g., video tagged with the same username as that of the person who uploaded the video). |
| D | | Irrelevant/Non Useful Tags |
| | 1 | Compound tag - Tags where words, phrases and sentences are joined together as one long text string. |
| | 2 | Multi Word Tags - Tags that as single words are meaningless, but placed in context with the other tags have meaning. (e.g. Celebrity name, Title of film, TV show, song, video game) |
| | 3 | Attention Attracting Tags – Tags that are assigned to attract attention to the video, that refer to popular search terms, but have no relevance to the video content. (e.g. Porn, Sex, Celebrity name.) |
| | 4 | Misspelling (e.g., 'Belguim' instead of 'Belgium') Whilst it may be obvious what the tag is meant to be, a misspelling obviously renders the tag useless in terms of subsequent users of the system who are searching for videos with that specific tag, unless they too misspell the tag/word. |
| | 5 | Unable to determine relationship Despite having attempted to look up either the meaning of the tag and whether the tag is a foreign word or not, tags which do not fit into any of the above categories will be deemed as unable to classify (e.g., nonsensical words). |
| | 6 | Foreign word/character |
| | 7 | Conjunctions and prepositions (e.g. the, in ,of, and) |

Findings

A large number of tags referred to people: some were famous people and some were people in the video, the creator, or the username of the uploader. 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. The majority, however, were names of people, bands or album titles that had been entered as two or more words. 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. Analysis of the dataset of 100 videos used for this research revealed that *'the'* constituted 1.4% of all tags and was also the most frequently used tag in the dataset.

Table 2 – Total number of tags and corresponding percentage of all tags, for each classification category.

| Classification Category | | No of tags | % age of all tags |
|-------------------------|---|------------|-------------------|
| A1 | Tag generically identifies what video is 'of' | 85 | 11.1% |
| A2 | Tag identifies video Category/Genre | 42 | 5.5% |
| 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% |
| C1 | Refining tag | 45 | 5.9% |
| C2 | Self-reference tag | 5 | 0.7% |
| C3 | Tag which explicitly denotes ownership of video | 8 | 1% |
| D1 | Compound 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% |

These findings suggest poor tagging practice of YouTube taggers and highlights that there is no shared vocabulary for tagging or a tagging standard as found in other systems like del.icio.us or Flickr (Ding et al. 2008). This is further highlighted by the lack of compound tags found in the dataset, only 0.4%. In contrast, Angus et al. (2008) found 12% of the tags in the Flickr data sample to be compound tags. However, a possible reason for the large percentage of compound tags is that Flickr handles tags entered as multi-word by converting them to a compound tag. YouTube has no system in place to try and encourage useful tags either via suggestions like in del.icio.us, or converting the user inputted text into a more

usable style like in Flickr. These findings emphasise the continued vocabulary problem faced by all tagging systems (Furnas et al., 1987; Golder & Huberman, 2005).

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. Further 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; Cattuto, 2007). This is not possible with this data sample as only 6.6% of tags occur more than once.

Despite having manually removed the majority of non-standard English characters from the database, some foreign words using standard English characters were overlooked, 8.7% were found, the joint fourth most common tag category. In retrospect, if all foreign words had been left in the dataset a more realistic gauge of non-standard English tags in the YouTube system could have been discovered. This would have been a useful finding to indicate the international appeal of the YouTube website, the variety of content and to give weight to the concept that the YouTube system would benefit from collaborative tagging, as multi language tags can help make the videos cross language barriers and be available for viewing by a wider audience.

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. In a collaborative tagging environment taggers will reuse tags they think describe the content well or are useful to their purpose and a standard is created for tag vocabulary in the system i.e. truncating or compounding names to form one tag, rather than two individual, not so meaningful tags (e.g., russell-crowe, russellcrowe, russell, crowe). The tagging method in YouTube does not afford this behaviour. More multi word tags were identified than compound tags. Multi word tags may be meaningful when displayed with other associated tags, but not in isolation. This renders them less useful for search, or browsing through tags. Whilst compound tags can be significant if seen in a tag cloud and could be used to browse tags to find videos, they are not useful for search as users will enter either single or multi keyword searches. This creates the problem of how to accept and handle multi word tags in a tagging system.

The classification suggested that the majority of YouTube tags in this dataset were of a subordinate level. 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, 2005). It could be the case that the random sample did not collect many videos with similar content, explaining why there were such a high percentage of tags that only appeared once in the dataset. Assuming the result is reflective of the YouTube system as a whole, if collaborative tagging was introduced, would the percentage of tags that occur once be reduced as more users entered tags that described the content? 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.

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. These tags are useful to a minority of users, as the majority of YouTube users probably want to be entertained, rather than to use the system to find specific video contents.

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)*.
- Browne, P. & Smeaton, A. F. (2004). Video information retrieval using objects and ostensive relevance feedback. In *Proceedings of the 2004 ACM Symposium on Applied Computing* (Nicosia, Cyprus, March 14 - 17, 2004). SAC '04.
- Capra, R. G., Lee, C. A., Marchionini, G., Russell, T., Shah, C., & Stutzman, F. (2008). Selection and context scoping for digital video collections: an investigation of YouTube and blogs. In

- Proceedings of the 8th ACM/IEEE-CS Joint Conference on Digital Libraries (Pittsburgh PA, PA, USA, June 16 - 20, 2008). JCDL '08. ACM, New York, NY, 211-220.
- 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.
- Cha, M., Kwak, H., Rodriguez, P., Ahn, Y., & Moon, S. (2007). I tube, you tube, everybody tubes: analyzing the world's largest user generated content video system. In Proceedings of the 7th ACM SIGCOMM Conference on internet Measurement (San Diego, California, USA, October 24 - 26, 2007)
- 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)
- Gill, P., Arlitt, M., Li, Z., & Mahanti, A. (2007). Youtube traffic characterization: a view from the edge. In Proceedings of the 7th ACM SIGCOMM Conference on internet Measurement (San Diego, California, USA, October 24 - 26, 2007). IMC '07. ACM, New York, NY, (pp.15-28).
- Golder, S. & Huberman, B. (2005) The Structure of Collaborative Tagging Systems [Online]. Available from:<http://arxiv.org/abs/cs.DL/0508082> Retrieved March 9, 2007
- 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
- Lambiotte, R. & Ausloos, M. (2005) Collaborative tagging as a tripartite network [Online]. Available from:<http://arxiv.org/abs/cs.DS/0512090> Retrieved March 11, 2007
- Li, X., Guo, L., & Zhao, Y. E. (2008). Tag-based social interest discovery. In Proceeding of the 17th international Conference on World Wide Web (Beijing, China, April 21 - 25, 2008).
- 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).
- Mathes, A. (2004) Folksonomies - Cooperative Classification and Communication Through Shared Metadata. [Online]. Available from: <http://www.adammathes.com/academic/computer-mediated-communication/folksonomies.pdf>. Retrieved May 18, 2007
- Thom-Santelli, J., Muller, M. J., And Millen, D. R. (2008). Social tagging roles: publishers, evangelists, leaders. In Proceeding of the Twenty-Sixth Annual SIGCHI Conference on Human Factors in Computing Systems (pp. 1041-1044).
- 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.
- Paolillo, J. C. & Penumarthy, S. (2007) The Social Structure of Tagging Internet Video on del.icio.us. System Sciences, 2007. HICSS 2007. 40th Annual Hawaii International Conference on, (pp. 85-85).
- Santos-Neto, E., Ripeanu, M. & Iamnitchi, A. (2007) The Organization of User Interests in Tagging Communities. In Proceedings of the International ACM/IEEE Workshop on Contextualized Attention Metadata: Personalized Access to Digital Resources. Vancouver, BC, Canada, June 2007 vol-266, (pp.11-18).
- Sen, S., Lam, S., Rashid, A., Cosley, D., Frankowski, D., Osterhouse, J., Harper, M. & Riedl, J. (2006) tagging, communities, vocabulary, evolution. CSCW '06: Proceedings of the 2006 20th anniversary conference on Computer supported cooperative work, (pp. 181-190).
- Shamma, D. A., Shaw, R., Shafton, P. L., & Liu, Y. (2007). Watch what I watch: using community activity to understand content. In Proceedings of the international Workshop on Workshop on Multimedia information Retrieval (Augsburg, Bavaria, Germany, September 24 - 29, 2007).
- Sinha, R. (2005) A cognitive analysis of tagging. Retrieved March 7, available from http://www.rashmisinha.com/archives/05_09/tagging-cognitive.html

- Trant, J. (2006) Social Classification and Folksonomy in Art Museums: early data from the steve.museum tagger prototype. ASIST SIG-CR workshop on Social Classification.
- Vander Wal, T. (2007) Folksonomy. Retrieved March 7, 2008. Available from:
<http://vanderwal.net/folksonomy.html>
- Voss, J. (2007) Tagging, Folksonomy & Co - Renaissance of Manual Indexing? In 10th international Symposium for Information Science (Cologne, Germany). Retrieved March 9, 2007
DOI= http://arxiv.org/PS_cache/cs/pdf/0701/0701072v2.pdf
- 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).
- Zollers, A. (2007) Emerging Motivations for Tagging: Expression, Performance, and Activism. Workshop: Tagging and Metadata for Social Information Organization. Proceedings of the Sixteenth International World Wide Web Conference (WWW '07), May 8–12, 2007, Banff, Canada.