Tree, funny, to_read, google: What are Tags Supposed to Achieve?

A Comparative Analysis of User Keywords for Different Digital Resource Types

Markus Heckner
Media Educational Science
University of Regensburg, Germany
markus.heckner@paedagogik.uni-regensburg.de

Tanja Neubauer
Information Science
University of Regensburg, Germany
tanja.neubauer@googlemail.com

Christian Wolff
Information Science
University of Regensburg, Germany
christian.wolff@sprachlit.uni-regensburg.de

ABSTRACT

Social tagging systems have become increasingly popular over the past years. Users' tagging practices have been little studied and understood so far. However, understanding tagging behaviour can contribute towards a thorough understanding of the tagging phenomenon from multiple perspectives. In the present paper, results of a comparative analysis of tag characteristics on the tagging platforms *connotea.org* (scientific articles), *del.icio.us* (bookmarks), *flickr.com* (photos), and *youtube.com* (videos) are presented. Results show that differences in tagging behaviour can be observed for different digital resource types. Finally, a short discussion of the possible implications of the results for the design of future tagging systems is presented.

Categories and Subject Descriptors

H.3.3 [**Information Storage and Retrieval**]: Information Search and Retrieval – *Search process*

General Terms

Design, Documentation, Experimentation, Human Factors

Keywords

Social tagging, tag functions, tag category model, folksonomy

1. INTRODUCTION

Social tagging systems allow users to organize, annotate and share content on the World Wide Web. Using these services, users may create a collection of items of personal interest and assign individual keywords (i.e. tags) to each of the resources in that collection. These user-created metadata can serve as additional document descriptors that can be matched against future search queries. Thus they offer an additional, user-driven layer of information description. The list of resource types that can be described by users currently includes browser bookmarks, photos, scientific papers, videos or even people ([1], [2]). Additionally, dedicated tagging systems have been introduced for knowledge management in corporate scenarios [3]. Although tagging lacks essential properties of controlled vocabularies like synonymy control, homonym discrimination and hierarchical structure (cf. [4]), tagging systems do provide many potential benefits for information retrieval: Tags are the manifestation of the users' language and serve as document descriptors for other users' search queries. Additionally, for social sharing of content as in media platforms like *Flickr* or *Youtube*, tagging currently appears to be the only possibility of semantic content description.

In order to fully utilize the potential of tagging for IR it is helpful to understand the range of possible functions a tag can have and whether these functions differ for different resource types. A better understanding of tagging practices is needed in order to design successful tagging platforms. However, little is understood about how users tag in practice: Do users *merely describe* document content or do they express personal relations or attitudes towards the resources by adding tags like *interesting* or *to_print*? Do users basically pick up tags from the content described or does their language usage significantly differ from media contents?

Some aspects of linguistic and functional properties of tags have been studied by Kipp and Campbell [5], Kipp [6], Golder and Huberman [7], Guy and Tonkin [8]. Heckner, Wolff and Mühlbacher [9] build on this previous work and present a functional and linguistic category model for tags used in the scientific bibliography management system *Connotea*. In the present paper, this model is used as basis for comparing the different systems and their associated resource types.

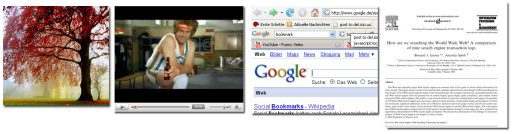


Figure 1 - tree, funny, google, toread – different resource types, different kinds of tags?

User-driven social tagging systems are a good example for the importance of (user) context in information interaction: Following the polyrepresentation hypothesis discussed in [10], we believe that social tagging systems offer an additional layer of information description that has the potential of better reflecting users' production and usage contexts for different media types. The polyrepresentation principle suggests that cognitively and func-

tionally different representations of retrieval objects might be used to enhance quality of IR results. Tagging is a form of polyrepresentation, since it adds another metadata layer to the representation of a digital resource within an IR system.

This paper is organized as follows: The next section reviews related work. Section 3 introduces our research design and methods for examining tag functions across four popular tagging systems (*Connotea*, *Del.icio.us*, *Flickr* and *Youtube*) as well as **TACS**, a php / MySQL tool which we have developed for the purpose of tag classification and analysis. The results of our empirical analysis of tag usage are presented in section 4. The paper concludes with a discussion of the implications of our results on the design of future tagging systems and an outlook on future research.

2. RELATED WORK

Several attempts have been made to interpret sets of user tags. Making sense of unstructured tags through clustering methods or similar mathematical approaches has been researched by Begelman et al. [11] and Halpin [12].

Kipp [6] presents a comparative analysis, examining non-subject related tag distributions across the systems *Del.icio.us*, *Connotea* and *CiteULike*[1]. However, Kipp's analyses are based on preselected articles that were tagged "for time, task and emotion" and no attempt to generalise on the distribution of these terms across system boundaries has been made.

Marlow et al. [13] turn their focus away from the individual tag and towards system-related aspects and "offer a model of tagging systems, specifically in the context of web-based systems, to help […] illustrate the possible benefits of these tools". Tonkin et al. [14] present an attempt to structure tags by assigning elements of the Dublin Core metadata standard to a set of tags. Finally, Heckner, Mühlbacher and Wolff [9] present a category model that reflects functional as well as linguistic characteristics of social tags. However, due to the intellectual classification effort the dataset used in this study was rather small (1091 tags applied to 500 resources) and the domain was limited to scientific articles posted on *Connotea*. As it appears to be the most comprehensive tag categorisation model yet, we will briefly discuss it in the following.

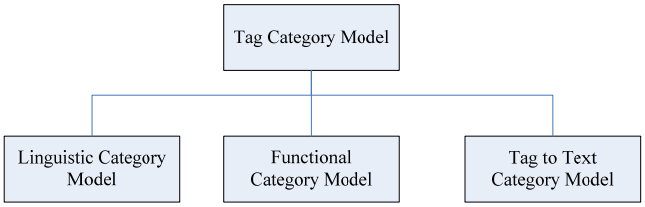


Figure 2 - Overview of tag category

Each of the subcategories (cf. Figure 2) of the overall category model represents an individual categorisation scheme for tags.

Functional categories into which a tag can fall are shown in Figure 3. The top-level distinction is made between tags that are directly related to the document, i.e. traditional metadata, and non-subject related tags. Subject related tags describe the resource by adding an author name, specifying a file type or referring to the source the document comes from, to name just some examples. The second subcategory of subject related tags concerns resource content: Tags can describe what a document is about or define a text category of the resource. Non-subject related tags on the other hand are in some way influenced by the users' current activities (time and task) or emotional state (affect). These tag types may be seen as indicators of usage and work context. Tag avoidance occurs when users deliberately decide not to add tags[2]. The linguistic model focuses on aspects like part of speech or number, while the tag to text model explores the possible relations between tag and text of the annotated resource.

3. RESEARCH QUESTIONS AND DESIGN

In the following, we present results from a comparative analysis and focus on the following research questions:

- Do intended tag *functions* differ for certain resource types? (E.g. do users of *Flickr* tend to tag more affectively than users of *Connotea*?)
- What is the relation between resource title and tags? Are there any significant differences in the percentage of tags which are taken from the *title of the tagged resource*?

We pick up the tag category system described in [9] and use it for selected tags from different types of tagging systems.

3.1 Selected Tagging Systems

For our comparative study, we chose to select popular tagging platforms for different media and informational entities: images, videos, bookmarks, and (scientific) documents. Our selection is based on an extensive list originally compiled by Shiri [15] to which a social dimension was added which distinguishes systems for private or recreational use from those which are mostly used in a working environment (cf. Table 1).

Table 1 - Overview and classification of existing tagging systems (selected systems in boldface)

Tagged items	Private	Professional context
photos	**flickr.com** bubbleshare.com	
bookmarks	**del.icio.us** blinklist.com spurl.net diigo.net taggly.com simpy.com shadows.com furl.net	
articles / news	slashdot.org	**connotea.org**, citeulike.com
people	myspace.com 43things.com	facebook.com
slides		slideshare.net slide.com
videos	**youtube.com** myvideo.de blip.tv	

[1] http://www.citeulike.org

[2] For a detailed description of the remaining two models see [9].

Connotea (scientific articles) is a scientific bibliography management system which enables researchers to tag the papers (and web resources) in their collection. *Del.icio.us* (bookmarks, web pages) is a social bookmarking management system where users can store and tag their favourite bookmarks online. *Flickr* (photos) is a photo sharing system which allows users to upload and tag pictures. *Youtube* (videos) is a system for video sharing that allows uploading and tagging of videos.

The four systems were selected because of their popularity which ensures availability of large datasets as well as their focus on a special type of digital resource. Each site is representative for one of the studied digital resource types and represents a prototypical instance of the respective resource category.

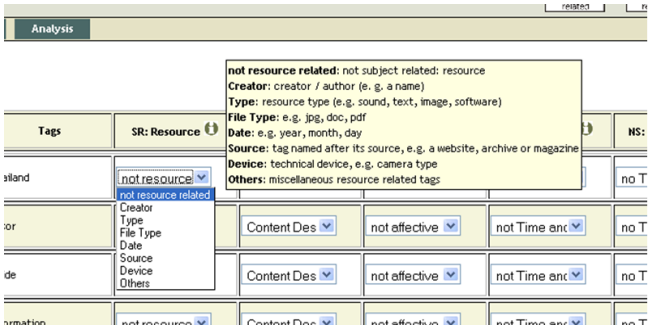


Figure 4 – TACS GUI: Classification in the functional model

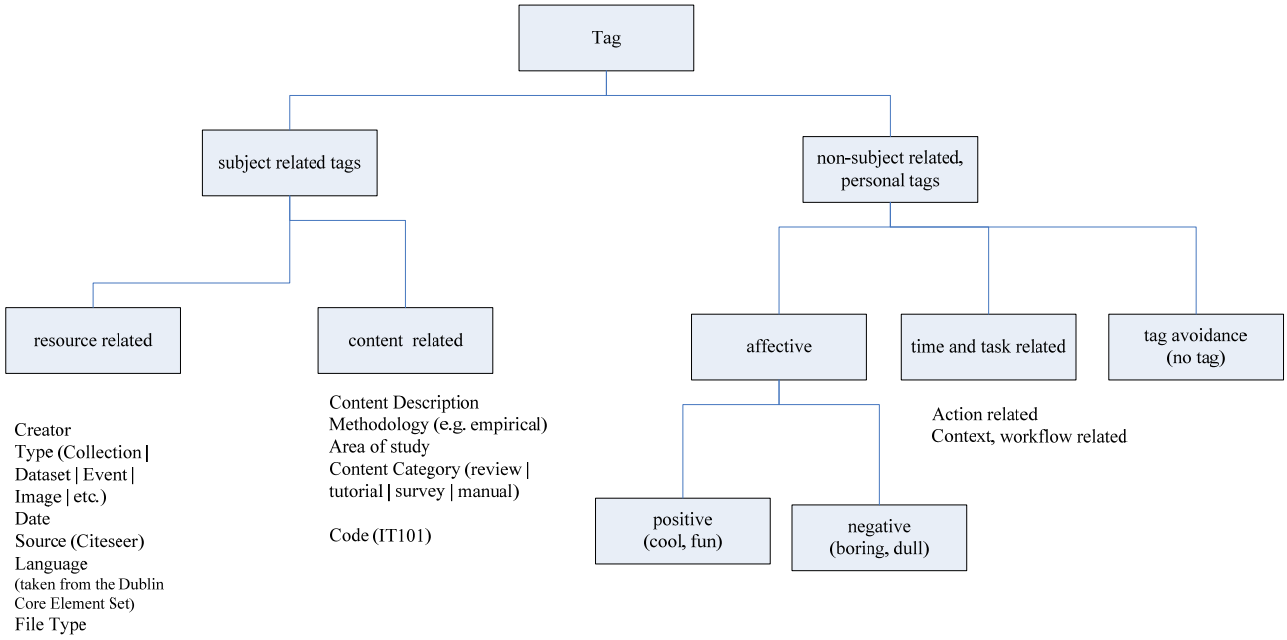


Figure 3 - Functional tag category model [9]

3.2 TACS - a Tool for Tag Annotation

In a previous study, we used MS Excel as a tool for raw data storage and for tag categorization. However, data entry and analysis using Excel turned out to be a lengthy, tedious and error-prone process not suited for collaborative annotation by several annotators. Consequently, a more systematic approach was called for. For this purpose, TACS (**T**ag **A**nalysis and **C**lassification **S**ystem), a tool for tag classification has been developed. TACS has been implemented in PHP and a MySQL database is used for structured relational data storage. TACS allows for tag categorization along the categories defined in [9], which were mapped to an appropriate MySQL data structure customized to fit the demands of the different tagging platforms.

Results can be exported to Excel or other tools for analysis. Tagging data is integrated into the database via TACS' web based import function. For data import, users select the source system and are subsequently prompted to choose the appropriate XML file from the computer's hard disk. To pull the information from the XML files, four different XML readers have been created to accommodate the individual structure of the files. The modular architecture of TACS allows for implementing additional XML readers to integrate further XML formats from other tagging platforms. The files are parsed and the tags together with various metadata elements like author or date are integrated into our MySQL data scheme.

3.3 Data Collection and Dataset

Data acquisition is a straightforward process, since the systems either provide web APIs or RSS feeds which can automatically be pulled periodically with a combination of Perl scripts and the command line web download tool *wget*. Data was gathered by downloading new posts on the respective systems every hour over

a period of several weeks. To avoid biases in the dataset, our sample subset was randomly selected from this very large dataset. For all four systems, the result format is XML and the files were imported in our categorization and analysis software described above. Our sample includes a total of 4012 tags taken from *Connotea* (1000), *Del.icio.us* (1010), *Flickr* (1001) and *Youtube* (1000). Table 2 gives an overview of tag numbers and associated documents (= resources).

Table 2 - Tags and resources

System	Tags	Resources
Connotea (articles)	1000	237
Del.icio.us (bookmarks)	1010	359
Flickr (photos)	1001	359
Youtube(videos)	1000	208
Overall	4012	1163

3.4 Categorization method

Tag categorization was carried out simultaneously by three annotators (students of information science). They were instructed to mark unclear cases and to leave comments in the comment box for these tags. In several consecutive workshops, which were also attended by the authors of the paper, these cases were discussed and the unclear cases were resolved. For cases which could not be sorted out in the workshops, a rule was in place to assign the tag to the content description category. For example, without getting hold of the users, it is not clearly determinable, whether the tag "web_service" is intended as content tag or used to organize a user's tasks.

4. RESULTS

We begin by presenting general findings, followed by comparing the system on a functional level. Finally the relation between the resource titles and the tags is explored. Discussion and interpretation follow in the subsequent section.

4.1 Average number of tags per resource

In order to examine tag numbers across the individual systems a single factor analysis of variance (ANOVA) was performed on the dependent variable number of tags per resource. The independent variable is either the type of tagging system or the resource type. The means and standard deviations are presented in Table 3. The analysis of variance reveals a significant difference, $F(3, 1159) = 18.38$, $p = .00$.

Table 3 - Tags per resource

	Connotea (articles)	Del.icio.us (bookmarks)	Flickr (photos)	Youtube (videos)
M	4.22	2.82	2.79	4.81
SD	5.10	2.23	4.02	4.18

Comparisons using the Tukey (HSD) test revealed that *Connotea* has a significantly larger number of tags per resource than both *Del.icio.us* (Cohen's d, .36) and *Flickr* (Cohen's d, .31). *Youtube* differs significantly from *Flickr* (Cohen's d, .49) and from *Del.icio.us* (Cohen's d, .59). *Del.icio.us* and *Flickr* as well as *Connotea* and *Youtube* do not differ significantly.

4.2 Words per tag

We also counted the number of words per tag in the four systems (see Table 4).

Table 4 - Words per tag

	Connotea.org (articles)	*Del.icio.us* (bookmarks)	flickr.com (photos)	youtube.com (videos)
M	1.27	1.11	1.40	1.01
SD	.64	.37	.77	.11

The analysis of variance for words per tag also reveals a significant difference, $F(3, 3723) = 94.35$, $p = .00$. A Tukey (HSD) post hoc test reveals a significant difference between all four systems. Effect sizes (Cohen's d) are: *Connotea – Del.icio.us* (.31), *Connotea – Flickr* (.18), *Connotea – Youtube* (.57), *Del.icio.us – Flickr* (.49), *Del.icio.us – Youtube* (.37), *Flickr – Youtube* (.71).

4.3 Subject vs. non-subject related tags

Figure 5 shows the distribution of subject vs. non-subject related tags across the four tagging systems.

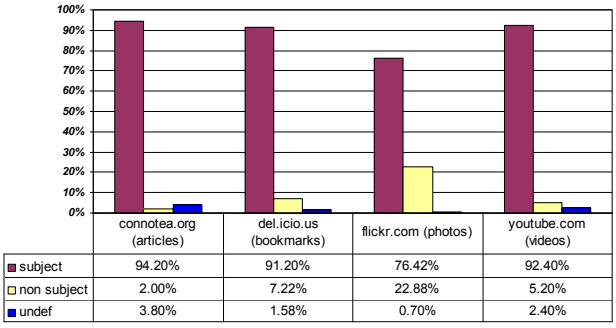


Figure 5 - Subject vs. non-subject related tags (tag avoidance included)

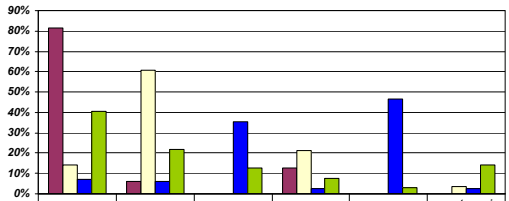
4.3.1 Tags Describing Resource and Content

Table 5 gives an overview of the distribution of subject related tags. *Connotea* and *Del.icio.us* both have lower numbers of resource related tags, while *Flickr* and *Youtube* have more resource descriptive tags.

Table 5 - Resource vs. content related tags

	Connotea.org (articles)	Del.icio.us (bookmarks)	Flickr.com (photos)	Youtube.com (videos)
Resource	16 (1.70%)	28 (3.04%)	82 (10.72%)	65 (7.03%)
Content	926 (98.30%)	894 (96.96%)	683 (89.28%)	859 (92.97%)
Overall	942 (100.00%)	922 (100.00%)	765 (100.00%)	924 (100.00%)

For all four systems and resource types the vast majority of tags (between 89% and 96%) describe the content of the resource. Subject related tags were categorized as *creator, type, date, source* and *device. Creator* tags refer to the creator of the resource (i.e. author, photographer or film producer).



	creator	type	date	source	device	uncategorized
connotea.org (articles)	81.25%	6.25%	0%	12.50%	0%	0%
del.icio.us (bookmarks)	14.29%	60.71%	0%	21.43%	0%	4%
flickr.com (photos)	7.32%	6.10%	35%	2.44%	46%	2%
youtube.com (videos)	40.63%	21.88%	13%	7.81%	3%	14%

Figure 6 – Distribution of resource related tags

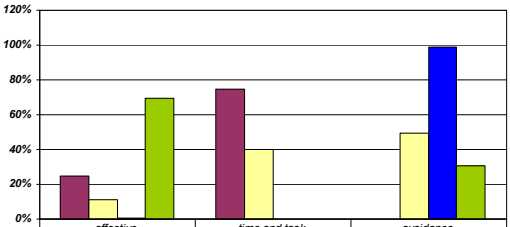
Type tags refer to the type of the resource (e.g. image, text, sound, software, video), *date* tags specify the creation date and *source* tags describe where the resource comes from (e.g. archive, website, magazine). *Device* tags could only be discovered in *Flickr* and *Youtube* and specify the camera which was used to produce the photo or video. *Youtube* also has a high percentage of type tags (more specifically: *video*). Numbers reported are relative frequencies. *Connotea* articles received most creator category tags. *Flickr* photos received most date and device tags, while many source and type tags were assigned to *Del.icio.us* resources (further details cf. Figure 6).

Table 6 - Distribution of content related tags

	***Connotea* (articles)**	***Del.icio.us* (bookmarks)**	***Flickr* (photos)**	***Youtube* (videos)**
language	0% (0)	0.22% (2)	0% (0)	0.12% (1)
content category	2.59% (24)	5.61% (50)	1.02% (7)	0.58% (5)
code	1.51% (14)	0.56% (5)	1.76% (12)	2.44% (21)
content description	95.79% (887)	93.39% (833)	97.07% (663)	96.62% (830)
uncategorized	0.11% (1)	0.22% (2)	0.15% (1)	0.23% (2)
Overall	100% (926)	100% (892)	100% (683)	100% (859)

4.3.2 Tags for Time and Task, Avoidance Tags

Only 20 *Connotea* tags, 73 *Del.icio.us* tags, 229 *Flickr* tags and 52 *Youtube* tags fell under the category of non-subject related tags. Relative frequencies and the distribution over the categories affective, time and task, and tag avoidance can be observed in Figure 7.



	affective	time and task	avoidance
connotea.org (articles)	25.00%	75.00%	0.00%
del.icio.us (bookmarks)	10.96%	39.73%	49.32%
flickr.com (photos)	0.87%	0.00%	98.69%
youtube.com (videos)	69.23%	0.00%	30.77%

Figure 7 - Distribution of non-subject related tags across systems

The following two listings contain all time and task and affective tags in our dataset.

Time and task related tags

Connotea *- course planning (2), createablog, createblogs, EXPLAIN, finding content, functional sites, irrelevance, missing link, NF-??, objectives, repository (2),startablog, Startblogging*; ***Del.icio.us*** *- AREA, bookmarks, buy, codetoread, getinvolved, help, imported (3), MyResearch (2), Office, readsoon_0710 (2), reference (5), research (3), sample, search, travel, vacation, zz.institution, zz.news*; ***Youtube*** **and** ***Flickr*** - (no time and task relation)

Affective tags (no correction of typing errors)

Connotea *- Best, lastminute, Mecca, meccabingo*[3]*, perfect*; ***Del.icio.us*** *- designsilike, exceptional, funny (3), geek, top;* ***Flickr*** *- sexy, sof;t;* ***Youtube*** *- "amaizng", amezing, :), condemned, cool, crazy, cute, fuck, funny (10), great, hey, hilarious, hittin, lame, nice, Oi!, retards, rich, sexy, sick, stupid, Stupid, sweet, weird, wow, yo, yuck*

Additionally, all affective tags were categorized according to their "polarity" as positive, neutral or negative. With the exception of *Flickr,* the clear majority of tags express a clearly positive view; at the same time, *only Del.icio.us* and *Youtube* have negative tags at all.

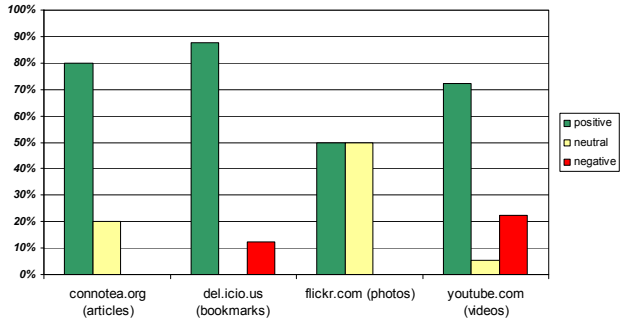


Figure 8 - Polarity of affective tags

Tag avoidance, i.e. the deliberate decision *not to tag* occurs in all systems with the exception of *Connotea*. For *Del.icio.us* 36 different avoidance tags are counted, 226 for *Flickr* and 16 for *Youtube*. *Flickr* users are avoiding tags much more often than users of other systems (cf. Figure 7). Examples of tag avoidance are given in Table 7.

Table 7 - Tag avoidance (examples)

Connotea	no tag avoidance occurred
Delicious	**, , - , ::**
Flickr	***???***
Youtube	***-, …., ..., dfgsdghsh, none, null*** *(2),*

4.4 Relation of tags to resource title

For the analysis of the tags' relation to the title of the resource all tag avoidance tags were excluded. Posting with titles is not man-

[3] *Mecca* and *meccabingo* describe a site for the game of bingo without any religious connotations and can thus be regarded as expression of positive affection.

datory in *Connotea* and *Flickr*. Consequently our dataset also includes resources without titles. These tags (214 *Connotea*, 17 *Flickr*) were also removed for the analysis of the relation of tags to title. The distribution of these remaining tags across the different relation categories can be observed in Figure 9 and Table 8.

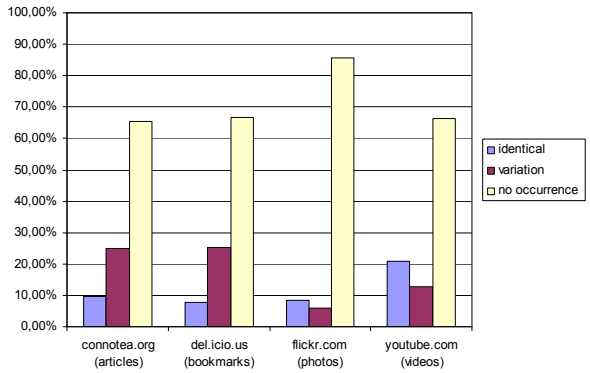


Figure 9 - Relation of tags to resource title across systems

Tags that are taken directly from the resource title are roughly evenly distributed between *Connotea*, *Del.icio.us* and *Flickr* (cf. Figure 9). *Youtube* stands out in this category: 21.03% of tags are identical to a word in the title.

Table 8 - Relation of tag to resource title

	***Connotea* (articles)**	***Del.icio.us* (bookmarks)**	***Flickr* (photos)**	***Youtube* (videos)**
identical	7.70% (77)	7.88% (77)	8.25% (64)	21.03% (209)
variation	19.6% (196)	25.38% (248)	5.80% (45)	12.68% (126)
no occurrence	72.70% (727)	66.73% (652)	85.95% (667)	66.30% (659)
Overall	100% (1000)	100% (977)	100% (776)	100% (994)

5. DISCUSSION OF THE RESULTS

Connotea resources (articles, M = 4.22) and *Youtube* resources (videos, M = 4.81) receive significantly more tags than *Flickr* resources (photos, M = 2.79) and *Del.icio.us* resources (bookmarks, $M = 2.82$).

For *Youtube* this can be attributed to a misconception of tagging by the users who seem to mistake the tag entry field for a text field for general free text description. The following examples list the title first, followed by the number of words in brackets and a comma separated list of tags.

corte de Luz *(4) - ELI, TITO, AND, ME;* ***for Taylor*** *(10) - mayday, parade, jilian, bowling, for, soup, 1985, lame, singing, dancing;* ***Poly Parrot Cheer*** *(6) - Poly, Drill, Team, At, Football, Game;* ***My Doppelgänger from the Future*** *(15) - future, world, of, warcraft, wow, doppelgänger, punch, me, in, the, face, please, just, do, it;* ***Fine Eateries at nowTV Channel 502*** *(7) - Good, dining, places, can, be, found, there;*

High numbers of tags per resource also occur when users copy and paste the title into the tag field. Apart from "complete" copy and paste modifications exist. Users copy but additionally alter the title information. *Youtube* treats words separated by spaces as distinct tags which leads to a tag chain in which each element constitutes one part of a whole phrase or sentence:

OK Go - Do What You Want *(7) - OK, Go, Do, What, You, Want, video;* ***Vie tickling baby Kenyon (8)*** *- Vie, tickle, baby, Kenyon, in, front, of, webcam;*

Another phenomenon we discovered was "overtagging" or very extensive tagging which occurs when users want to make sure their video is discovered by as many people as possible. Here a distinction between personal information management and tagging for the public becomes obvious: We assume that users who tag videos do not want to organize their personal collection, but rather want the video to be retrieved and viewed by as many people as possible. The following list contains examples of *overtagging* on *Youtube* (again, resource title in boldface):

Our daily bread *(15) - Worship, Preaching, Teaching, Disciple, Discipleship, Shepherd, Servant, Sheep, Bread, Bible, Jesus, Believer, Living, Spirit, Fruit;* ***Signs as done by Tesla Cover by The Localz*** *(14) - thelocalz, localz, locals, tesla, signs, five, man, electrical, washington, nc, north, carolina, david, winstead;* ***Interactive map tutorial*** *(11) - Forestry, environment, logging, nonprofit, foundations, Minnesota, Sweden, Finland, state, government, county;* ***BIASCA THE WITCH "V. Tognola interviews a priest from 1513"*** *(13) - biasca, ticino, frama, films, switzerland, victor, tognola, the, witch, interview, priest, church, documentary;* ***Broken Heart Part 1*** *(10) - Broken, Heart, Part, maple, story, babyprincexdj, skispark, Montana, friends, :);* ***Story [Ruka && Michi amv]*** *(9) - sailor, moon, michiru, haruka, uranus, neptune, air, anime, manga;* ***Golf Can Be A Simple Game - Ebook Information*** *(9) - golf, swing, ebook, improvement, tips, putting, training, practice, workout;*

Examples for extensive tagging could also be discovered for *Flickr*:

--- no title --- *(9) - dexter, puppy, dog, pet, cavalierkingcharles, cavalier, king, charles, spaniel;* ***Tom Hayden*** *(16) - barack obama, cia, democratic debate, election08, fbi, hillary, hillary clinton, john edwards, kodak theater, lapd, liveblog, liveblogging, obama, security, tsa, wolf blitzer;* ***20080121-DSC_5612*** *(16) - afterparty, bunny, club, dj, film festival, night club, nightlife, park city, party, rabbit in the moon, rave, ritm, star bar, sundance, sundance film festival, utah;*

Noteworthy from an Information Retrieval point of view is that some users include synonyms, near synonyms and spelling variations in their tag descriptions: *training, practice, workout; nc, north carolina; thelocalz, localz, locals; anime, manga.*

When comparing words per tag, all systems differ significantly, *Youtube* being the system with the lowest number of words per tag (M = 1.01). This can be explained by *Youtube*'s word segmentation algorithm: Spaces are counted as tag separators and thus tags which were intended as one descriptor by the tagger (e.g. *Information Retrieval*) are split up into the two distinct tags *information* and *retrieval*. *Connotea* on the other hand (the system with the highest number of words per tag (M = 1.27) also uses spaces as default delimiters but allows users to explicitly specify multiword tags by enclosing them in double quotes (e.g. *"Information Retrieval"*). *Del.icio.us* does not allow multiword tagging, so that assigning a descriptor with two words is only possible by using *CamelCase*, underscores or similar circumvention strategies.

On average, *Del.icio.us* and *Flickr* users are more reluctant in the number of tags they assign. The low number of *Flickr* tags might

again be attributed to the users have: Photos can be shared and shown to other users, when they are pointed to the right album via URL, retrieval is not critical, since the items in the album can easily be browsed and photos are instantly self descriptive, when viewed by a user.

5.1 Resource and Content related Tags

The tendency of *Flickr* towards resource (or context) related tags can be explained by the fact that many users tag with the date the photo was shot (29) or describe the camera they used (37). Examples for these device category tags are: *Canon 1D Mark II N, Canon 17-40mm f/4 L, canon eos 250d, nikon, lomo, fuji, Nikon D300.* For *Youtube* only two device tags (*mobile, cellphone*) are present. Device tags are not used at all in the other systems.

A reverse trend between *Flickr* and *Youtube* can be observed: *Flickr* tags tend to describe the location the photo was shot (173) and the person that is in the photo (41). These numbers are almost directly inverse in *Youtube*, where only 44 tags describe the filming location and 140 tags describe the persons in the video.

5.2 Time, Task, Affect, and Tag Avoidance

The idea of time and task related tags was first brought up by Kipp (2006), who found that over 16% of all *Del.icio.us* tags could be categorized as "related to a users current projects or activities". However we could not confirm these result in a previous study [9] which examined *Connotea* tags: Only about 2% of *Connotea* tags in our dataset could be categorized as time and task related. The present study cannot confirm Kipp's initial results as well: The system with the highest percentage of time and task related tags is *Del.icio.us* (2.9%), followed by *Connotea* (1.5%), *Flickr* (0.1%) and *Youtube* (no time and task related tags at all).

Affective tags were also first studied by Kipp (2007) and described as "dwelling on a user's emotional response to a document". However, she does not report actual frequency data but focuses on the role of affective tags in the tagging process. *Youtube* is the system with the highest number of affective tags (36 or about 3.6%).

However, affect is not expressed as frequently as we had originally assumed to be the case at least for leisure-oriented systems like *Flickr* or *Youtube*. In all four systems users show the tendency to express positive emotional responses rather than negative ones.

Tag avoidance is very popular with *Flickr* users, about 25% (226) of all *Flickr* tags were classified as tag avoidance. Interpretation is difficult to achieve without getting hold of the users and their intentions. A possible explanation might be an additional organization principle offered by *Flickr*: Users can create photo albums (i.e. sets of images) to which pictures can be added. Consequently, *Flickr* photos often come in packs of pre-organized containers. Users might not feel the need to explicitly tag each photo as "*holiday, summer, 2007*" but instead drop the content into the surrounding container which is labelled as a whole. Furthermore, *Connotea*, *Del.icio.us* and *Youtube* lack this property of formal item sets and solely rely on tags as organization principles.

5.3 General trends

Our initial hypothesis that different types of resources are tagged functionally different could be confirmed:

- photos are tagged for *content*
- photos are tagged for *location*
- photos are often *untagged*
- photos are tagged with the *camera device name*
- videos as well as photos are often tagged *extensively*
- videos are tagged for *persons*
- scientific articles are tagged for *time* and *task*

However, we have to be cautious when interpreting these results: Users differ on social and task dimensions which have implications for both their understanding of tagging and their goals. While uploading an item on *Youtube* clearly entails a user's motivation to make the item accessible to other users, this is not so clear for posting an item to *Connotea* or *Del.icio.us* where aspects of personal information management (PIM, [16]) come into play. Answering these questions is not possible with randomly chosen samples of tags from unknown users. Future research depends on directly gathering data from users on their tasks, motivations, and goals.

5.4 Design implications for tagging systems

Tagging is a form of explicit feedback which could be used for filtering search results in a social tagging system. Affective or time and task related information might be a valuable source for collaborative filtering techniques. The success of Amazon's *"... you might also like what people with similar interests as you have also bought..."* approach to recommend items based on implicit user feedback proves the feasibility of this approach (cf. Linden [17]). Desiring an item from Amazon and buying it seems to be a logical and causal sequence of steps. The tagging process lacks this causality: The unstructured and "free" nature of tagging does not contribute much to eliciting this kind of feedback from users. Yet, simply because users do not express their emotional response or opinion about a resource does not mean that it does not exist. Therefore the need for some form of prompting the users to explicitly state their opinion about a resource arises. Approaches like this can be found on *Youtube*, where users rate video quality. Another feasible attempt might be to include a "readability" rating for *Connotea* or a "funnyness" scale for *Youtube*, thus using tags as quality markers beyond content description. However, voluntary user participation in *Youtube* is rather low: Ratings only occur for just 0.22% of all views. Comments, which require even more active participation, account for a mere 0.16% of all views (Cha et al.[18]).

Results show that tendencies of users to tag photos and videos for content, location and persons exist. One step towards semantically enhanced tags which would allow for structured retrieval could be achieved by dividing the tag entry field into separate categories. Including "people and place fields" for *Flickr* tags might open the possibilities for more elaborate semantic queries. Information needs like "I would like to see all pictures or videos of *George **in** Washington*" are currently unlikely to be satisfactorily solved by *Flickr* or *Youtube*, but could be accommodated by structured tag entry fields, which could add meaning to a tag. For a further discussion of adding structure to tags by applying Semantic Web technology to tagging systems see Gruber [19].

6. CONCLUSION AND FURTHER RESEARCH

From our analysis of tagging data, several trends become visible: On the functional level, *Flickr* users show a greater tendency towards tag avoidance than users of other systems. Affective tags that express an emotional aspect could be discovered in all systems but on a moderate scale and show a tendency towards expressing positive rather than negative or neutral emotions.

Apart from the resource type other factors play an important role, and have to be explored further: What intention do users have? Do they want to organise, save or distribute? Does tagging have the potential of significantly going beyond content description? Influences of user interface design and functionality (tag suggestion algorithms, separation of multi word tags, etc.) on tagging behaviour remain an interesting question for further research.

Finally it has to be noted that tag suggestion algorithms and interfaces differ for the selected systems. This is a confounding variable that limits the strength of the presented results. Further studies in a controlled environment have to be carried out to level out these limitations.

7. REFERENCES

[1] Farrell, S. and Lau, T. Fringe Contacts: People-Tagging for the Enterprise. In *Proceedings of the Proceedings WWW2006, Collaborative Web Tagging Workshop* (Edinburgh, 2006). ACM Press, New York, NY.

[2] Voß, J. Tagging, Folksonomy & Co - Renaissance of Manual Indexing? In *Proceedings of the Open Innovation. Proc. 10th International Symposium for Information Science* (2007). UVK, Constance.

[3] Maßun, M. From Personal Information Management Towards Collaborative Information Management in Enterprises. In *Proceedings of the Open Innovation. Proc. 10th International Symposium for Information Science* (Cologne, May, 2007). UVK, Constance.

[4] Macgregor, G. and McCulloch, E. 2006. Collaborative tagging as a knowledge organisation and resource discovery tool. Library Review. 55 (2006), 291-300.

[5] Kipp, M. E. I. and Campbell, D. G. Patterns and Inconsistencies in Collaborative Tagging Systems: An Examination of Tagging Practices. In *Proceedings of the Proceedings of the 2006 Annual Meeting of the American Society for Information Science and Technology* (2006)

[6] Kipp, M. @toread and cool: Tagging for time, task and emotion. In *Proceedings of the Proceedings 8th Information Architecture Summit* (Las Vegas, Nevada, 2007)

[7] Golder, S. and Huberman, B. A. 2006. Usage patterns of collaborative tagging systems. Journal of Information Science. 32 (2006), 198-208.

[8] Guy, M. and Tonkin, E. 2006. Folksonomies - Tidying up tags? D-Lib Magazine. 12 (2006).

[9] Heckner, M., Mühlbacher, S. and Wolff, C. 2008. Tagging tagging. Analysing user keywords in scientific bibliography management systems. Journal of digital information (JODI). 9 (2008).

[10] Larsen, B., Ingwersen, P. and Kekäläinen, J. The polyrepresentation continuum in IR. In *Proceedings of the 1st international conference on Information interaction in context, IIiX 2006* (Copenhagen, 2006). ACM, New York, NY.

[11] Begelman, G., Keller, P. and Smadja, F. Automated Tag Clustering: Improving search and exploration in the tag space. In *Proceedings of the Proceedings WWW2006, Collaborative Web Tagging Workshop* (2006). ACM Press, New York, NY.

[12] Halpin, H., Robu, V. and Shepherd, H. The complex dynamics of collaborative tagging. In *Proceedings of the WWW '07: Proceedings of the 16th international conference on World Wide Web* (Banff, Alberta, 2007). ACM, New York, NY.

[13] Marlow, C., Naaman, M., Boyd, D. and Davis, M. HT06, tagging paper, taxonomy, Flickr, academic article, to read. In *Proceedings of the HYPERTEXT '06: Proceedings of the seventeenth conference on Hypertext and hypermedia* (Odense, 2006). ACM Press, New York, NY.

[14] Tonkin, E., Baptista, A. A., van Hooland, S., Resmini, A., Mendéz, E. and Neville, L. Kinds of Tags: a collaborative research study on tag usage and structure. In *Proceedings of the Paper presented at The 6th European Networked Knowledge Organization Systems (NKOS) Workshop Workshop at the 11th ECDL Conference* (Budapest, Hungary, 2007)

[15] Shiri, A. Trend Analysis in Social Tagging: An LIS Perspective. In *Proceedings of the Paper presented at The 6th European Networked Knowledge Organization Systems (NKOS) Workshop Workshop at the 11th ECDL Conference* (Budapest, 2007)

[16] Boardman, R. and Sasse, M. A. "Stuff goes into the computer and doesn't come out": a cross-tool study of personal information management. In *Proceedings of the CHI '04: Proceedings of the SIGCHI conference on Human factors in computing systems* (Vienna, 2004). ACM, New York, NY.

[17] Linden, G., Smith, B. and York, J. 2003. Amazon.com recommendations: item-to-item collaborative filtering. Internet Computing, IEEE. 7 (2003), 76--79.

[18] Cha, M., Kwak, H., Rodriguez, P., Ahn, Y.-Y. and Moon, S. I tube, you tube, everybody tubes: analyzing the world's largest user generated content video system. In *Proceedings of the IMC '07: Proceedings of the 7th ACM SIGCOMM conference on Internet measurement* (San Diego, CA, 2007). ACM, New York, NY.

[19] Gruber, T. 2007. Ontology of Folksonomy: A Mash-up of Apples and Oranges. International Journal on Semantic Web & Information. 3 (2007).