M. Herczeg & M. C. Kindsmüller (Hrsg.): Mensch & Computer 2008: Viel Mehr Interaktion.
München: Oldenbourg Verlag, 2008, S. 307-315

Between "Instructions" and "Diy": Tagging in Learning Communities

Milena Reichel, Heidi Schelhowe

Digitale Medien in der Bildung, TZI, Universität Bremen

Abstract

In this paper, we discuss the novel technology tagging and the results of analyzing a learning community in a popular system that relies on tagging namely YouTube. We present our findings that young people use tagging as social technology and tags to create a vocabulary among friends and communities. We argue that tags can be used to facilitate sharing and understanding in learning communities. Finally we give an example how to apply tagging to a learning environment.

1 Introduction

According to studies i.e. the JIM study (Medienpädagogischer Forschungsverbund Südwest 2007) children and young people use the Internet mainly as a social medium. Social Network Sites (SNS) for example *Facebook*[1] or *MySpace*[2] are very popular among youngsters. Not surprisingly the most important topic they relate digital media to is "love and friendship".

The popular Social Network Sites among other novel services and applications rely on some new technologies and usage patterns. User-created content, open "Application Programming Interfaces" (APIs) or *tagging* enjoy a huge success in usage rates on the Internet and are often summarized under the umbrella term "Web 2.0" (O'Reilly 2005).

While APIs allow other applications and therefore other developers to access data of a particular system, tagging enables the user herself to annotate data with metadata. We will go into more detail on tagging in the next section.

[1] http://www.facebook.com
[2] http://www.myspace.com

2 Tagging

As a definition of tagging one can say that tagging is the process of labeling items i.e. pictures or URLs with one or more free text keywords.

“Taggers” include children and young people as well as adults. Popular examples of systems that use tagging are *Delicious*[3], a bookmark sharing service; *Facebook*, the popular SNS we mentioned before; or *YouTube*[4], a service to share videos.

Every user tags his/her item for example pictures. The entity of all users’ tags in a system builds a navigation structure that is called “Folksonomy” (Mathes 2004) – short for "folks" and "taxonomy" because of its quality as a bottom-up organized, decentralized hierarchic structure. The Folksonomy typically serves as navigation structure for the website.

Current research on tagging systems includes usage patterns in tagging systems (Golder & Huberman 2005), data retrieval and advanced search (Begelman et al. 2006), alternative visualizations (Dubinko et al. 2007), but also user motivations and experiences related to tagging (Marlow et. al. 2006). We will focus on the latter in our study on YouTube.

3 Study on YouTube

YouTube is a well-known service to upload and share videos. Videos require both a tag and assignment to a given category during the upload process. An important aspect of YouTube related to (young people’s) learning is that it is well known and used among youngsters according to i.e. the JIM study.

In their results 9 % of the interviewees (aged between 12 and 19) use YouTube actively (produce their own content), 60 % use it passively, 17 % heard of the name while only 14 % never heard about the service before the interview.

To gain a general impression of contents, users and tags in the YouTube system we started by analyzing YouTube’s popular content. The data sample we took consists of the featured, the most popular and the most discussed videos during the time period between August 22nd 2007 and September 27th 2007. It was acquired by using YouTube's API and parsing YouTube’s RSS-feeds of most popular and most discussed videos hourly during the time period. We harvested video descriptions and user profiles. The sample included 13.868 users and videos with a total amount of 65.535 tags (28.364 distinct tags).

The average (specified) age of the users in the sample was 23.11; 68.2 % of the users indicated they were male. The most common tags in the data sample were: “2007” (with 0.6 %), “funny”, “video”, “the” and “music”.

[3] http://www.delicious.com
[4] http://www.youtube.com

An interesting insight was the difference in the different age groups concerning their use of social connection features (called friend or family in YouTube). Young people use the friendship function more often than the older users.

The popular tags suggest people use YouTube mainly for fun. There is no evidence that YouTube is used for learning. But within YouTube smaller communities based on shared interest can create their own spaces. The magazine “Make”[5] that publishes instructions on how to make technology projects for example has a community in YouTube. We will refer to it as “Make”.

Make had 1667 members (at 31st October 2007). The users are 33.81 years old in average (according to the age they submitted) and the percentage of indicated males is 74.5 %. The popular tags in the Make are quite different from YouTube's popular tags. Most common are: “make” (5.1 %), “diy”, “bre”, “instructions”, “howto” and “pettis”. The tags clearly center on learning how to make things.

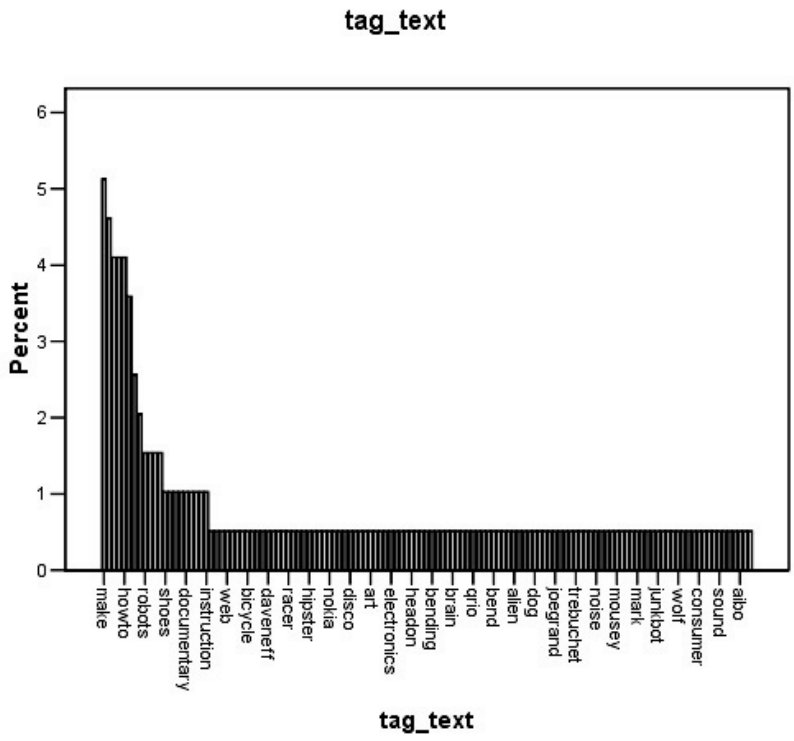

Figure 1: Most common tags of the Makers

[5] http://www.makezine.com

Some of the tags are understandable for an outsider of the community e.g. “make” and “diy”. “Pettis” and “bre” on the other hand are not. They are part of the special vocabulary the community developed and require the context knowledge that “Bre Pettis” is the producer of a weekly video show of the Make magazine to be meaningful.

Some of the Makers are connected in a stronger way by using the friendship functionality of YouTube. From all members that have friendships we harvested the user profiles of friends, leading to a total amount of 6397 users related to Make.

We further analyzed the resulting group by using “Social Network Analysis” (SNA). SNA is an own research field that we touch only briefly in respect to modeling users and tags of YouTube as a social network. SNA describes social groups for example local communities or organizations as networks, where nodes represent persons and edges between nodes relations of those people. For our SNA we modeled the users of YouTube as nodes and added a relation whenever two users used the same tag. We compared a group of befriended users with a group of random users.

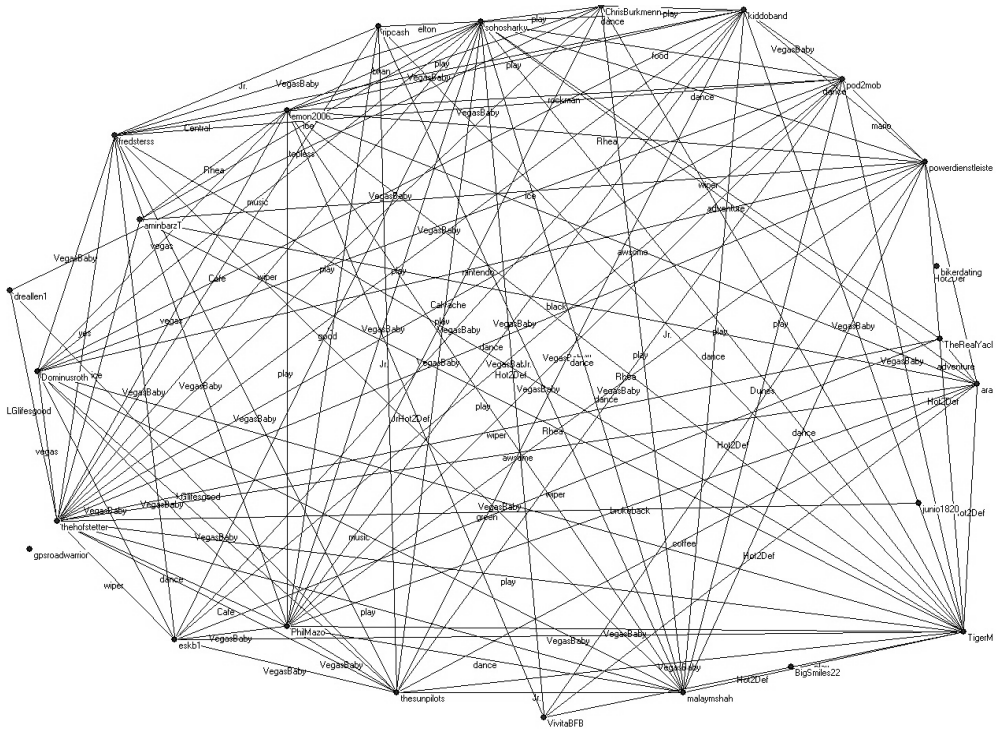

Figure 2: Network of friends

The first step of our analysis was to identify a group within the members of Make that was related through friendships in a dense way. To identify such strongly connected groups we went through the following steps:

- First we performed a network reduction and removed vertices with a degree smaller than one. In our model these are users that have no friendships.
- Afterwards we removed all multiple lines and loops.
- Within the resulting partition we searched for weakly connected components.
- In these components we identified k-cores.
- We extracted the cores with k=10.

The extracted vertices had a number of 41 in total.

As we wanted to find out whether there was a correlation between friendships and tag usage, we harvested all tags of the 41 strongly connected users and added an edge between them if they shared a tag. Within the resulting partition we searched for weakly connected components.

As a result we ended up with 24 users that were connected through their tag usage. You can see from the resulting graph (cf. Figure 2) that these users are mainly connected with each other; there are only two users, “BigSmiles22” and “bikerdating” that belong to weakly connected components but are kind of outsiders. The density of the network without loops is 0.4057971.

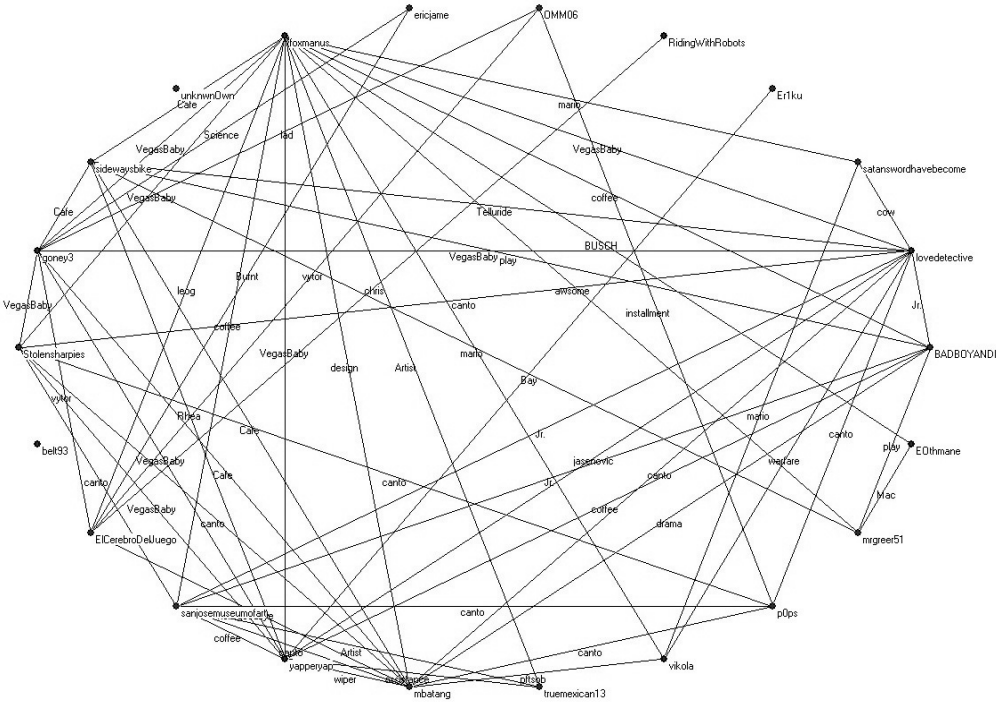

Figure 3: Network of random users

In the next step we created a network without the friendship connection for comparison. We chose a random set of the Makers and their friends in the same size (41) and reduced the network by deleting loops and multiple lines. The edges represent again tags used by a connected pair of users. 22 users of the set have a degree larger then 2. The density is lower namely = 0.2554113 compared to the network amongst friends.

When one compares both resulting graphs it is easy to see that befriended users are much closer connected through their tag usage than the random users of Make.

4 Results and Implications

Studies indicate that children and young people use the Internet mainly as a social medium. We showed that this focus can also be found in tagging systems, i.e. in YouTube where especially young people use social connections frequently, more often than adult users. We also showed that these connections and tagging vocabulary are interrelated. Users connected through friendship more often use the same tags than users that are part of a special interest community.

We also would like to point out that tagging systems even though they might center mainly on different topics allow communities to create their own informal learning spaces as we saw it within the Make community.

We conclude that tagging can be a helpful tool to design and create these spaces as the community shapes its environment itself. People interested in similar topics can find and join the communities through tags. Tagging and how vocabularies build is related to social connections among the users.

We argue that we can apply tagging to the design of learning environments in other domains e. g. in approaches that rely on Constructionism (cf. Papert 1991) to give the communities a technology to design their learning spaces and to build a social network even within the larger learning community. When we create tagging software for learning we should:

- Focus on social connections and allow friendships among different domains
- Use tagging to annotate data as well as other parts of the environment
- Allow the learners to develop their own vocabulary and shape their learning space by tagging

Especially in learning environments that apply a Constructionist approach where the main idea is to learn through designing a personally meaningful object, designing an environment oneself is valuable for the learning process.

5 Prototype

Based on the conclusions from the last section we implemented a programming environment called Amici that can be used to program one’s own Smart Textile project with our EduWear construction kit (Reichel 2007), Leah Buechley’s beautiful LilyPads (Buechley et al. 2008) or standard Arduino hardware (Mellis et al. 2006). Textiles and fashion are deeply embodied in young people’s culture therefore we consider them as suitable medium to learn programming.

In Amici we combine a visual block interface inspired by the building block metaphor (Begel 1999) with textual programming based on the Open Source hard- and software Arduino.

Tagging is used in different way to ease sharing of programs and projects on the one hand but also the whole programming process on the other hand:

5.1 Start by the Example

Instead of having one’s project as the center of the programming environment we start by showing the “world”: the set of users and projects that are related to the current user and his or her project. User and projects are related by a calculation of tag similarity or by using the friendship function (see Figure 4).

5.2 Tag Blocks

Each block represents a command. The usage of these commands might not be self-explaining to the users. In our observations of children using block-based software we found for example that using an “if” block when they intend to ask for values from a sensor is an unfamiliar concept. In Amici users can add tags to blocks. Children tagged the block with the “on” command with tags like “LED” or “motor”, “driving” and “blink” for example (see Figure 4). The tags users give to blocks are also used in the tag similarity calculation.

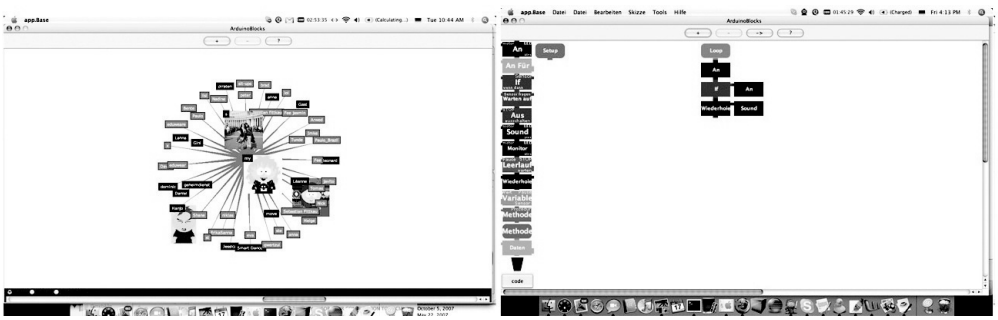

Figure 4: From left to right: start by example, tag blocks

5.3 Share and Publish Projects

Children and young people can publish their projects to a web platform in order to share them. Projects consist of code in form of blocks or source code, pictures, tags and further descriptions. They are saved within the web-based database and are accessible via the search menu (see Figure 5) or the browser-based client. After publishing a project it (and the corresponding user) will be included in calculation of similarity.

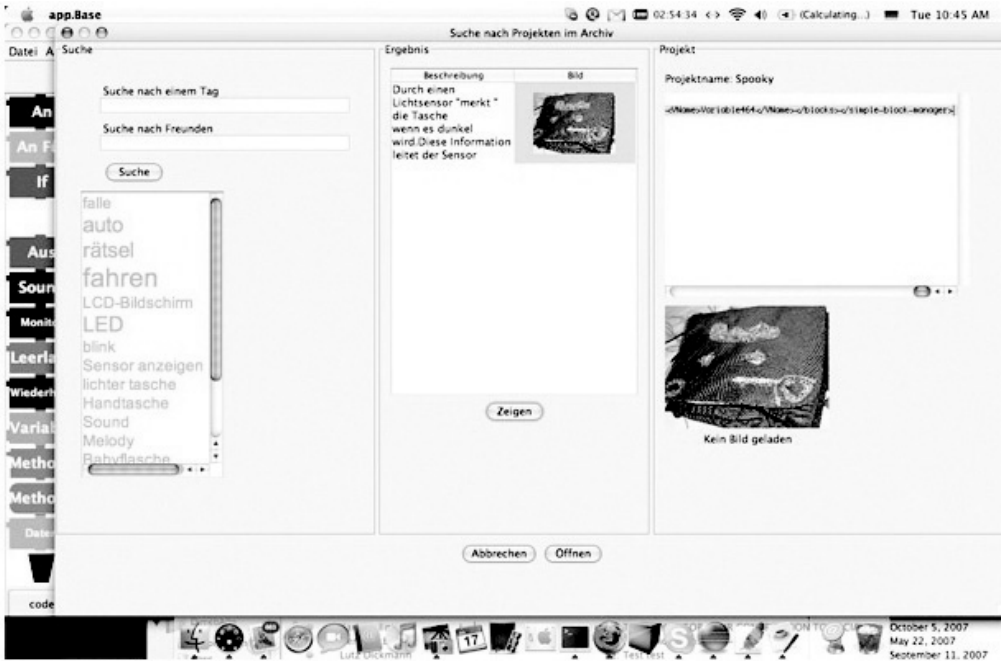

Figure 5: Share and publish projects

The similarity of projects and users we already mentioned is calculated by Begelmann et al.'s tag clustering algorithm (Begelmann et al. 2006) in an extended form. In the first step we calculate the similarity of each pair of tags by the probability that both tags are used for the same item. We also consider further information in the calculation e.g. what kind of sensors or actuators are used or whether the users of the projects are connected through friendship. The resulting graph is grouped into clusters by deleting the edges where the weight is small.

5.4 Social help system

As help system we recommend other examples based on similarity of the projects instead of confronting the user with a written form of manual.

6 Conclusion

Within the prototype we use tagging in different ways: to annotate elements of the programming language to allow user to create their own space, as basis to recommend projects as a context-sensitive social help and to annotate projects and publish them into the Internet and create a navigational structure based on the tags.

We will use our prototype in real learning scenarios and collect a larger set of tags. Afterwards we will compare the rather formal communities with the informal ones in YouTube. For the evaluation of the prototype we suggest to add a qualitative perspective to learn how users perceive the tagging process.

References

Begel, A. (1996). Logoblocks: A graphical programming language for interacting with the world. Master's thesis, S.B. Thesis, MIT Department of Electrical Engineering and Computer Science.

Begelman, G. & Keller, P., & Smadja, F. (2006). Automated tag clustering: Improving search and exploration in the tag space. *Collaborative Web Tagging Workshop at WWW2006*, Edinburgh.

Buechley, L., Eisenberg, M., Catchen, J., & Crockett, A. (2008). The lilypad arduino: Using computational textiles to investigate engagement, aesthetics, and diversity in computer science education. In: *Proc. SIGCHI conference on Human factors in computing systems*, New York, NY. ACM.

Dubinko, M., Kumar, R., Magnani, Novak, J., Raghavan, P., & Tomkins, A. (2007). Visualizing tags over time. *ACM Trans. Web*, *1*(2).

Golder, S. & Huberman, B. A. (2005). The structure of collaborative tagging systems. *Journal of the American Society for Information Science and Technology.*

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*, S. 31–40, New York, NY: ACM Press.

Mathes, A. (2004) *Folksonomies – cooperative classification and communication through shared metadata*, http://www.adammathes.com/academic/computer-mediated-communication/folksonomies.html, last checked 19.10. 2007.

Medienpädagogischer Forschungsverbund Südwest (2007). *JIM-Studie 2007*. http://www.mpfs.de, last checked 10.01.2008.

Mellis, D., Banzi, M., Cuartielles, D., & Igoe, T. (2007). Arduino: An open electronics prototyping platform. *alt.chi section of the CHI 2007 conference in San Jose* (CA).

O'Reilly, T. (2005). *What Is Web 2.0. Design Patterns and Business Models for the Next Generation of Software.* http://www.oreilly.com/pub/a/oreilly/tim/news/2005/09/30/what-is-web-20.html, last checked 19.10. 2007.

Papert, S. & Harel, I. (1991). Situating constructionism. In Seymour Papert, I. H., editor, *Constructionism*, Ablex Publishing: New York, NY.

Reichel, M. (2007). EduWear: Ein Construction Kit für Smarte Textilien und Wearable Computing. *Proc. Informatik 2007. Informatik trifft Logistik, Bremen.*