Of course we share! Testing Assumptions about Social Tagging Systems

A Case Study of the Social Tagging System BibSonomy

Stephan Doerfel
University of Kassel
Kassel, Germany
doerfel@cs.uni-kassel.de

Daniel Zoller
University of Würzburg
Würzburg, Germany
zoller@informatik.uni-wuerzburg.de

Philipp Singer
Technical University Graz
Graz, Austria
philipp.singer@tugraz.at

Thomas Niebler
University of Würzburg
Würzburg, Germany
niebler@informatik.uni-wuerzburg.de

Andreas Hotho
University of Würzburg
Würzburg, Germany
hotho@informatik.uni-wuerzburg.de

Markus Strohmaier
University of Koblenz-Landau and GESIS
Koblenz and Cologne, Germany
strohmaier@uni-koblenz.de

ABSTRACT

Social tagging systems have established themselves as an important part in today's web and have attracted the interest from our research community in a variety of investigations. The overall vision of our community is that simply through interactions with the system, i.e., through tagging and sharing of resources, users would contribute to building useful semantic structures as well as resource indexes using uncontrolled vocabulary not only due to the easy-to-use mechanics. Henceforth, a variety of assumptions about social tagging systems have emerged, yet testing them has been difficult due to the absence of suitable data. In this work we thoroughly investigate three available assumptions – e.g., is a tagging system really social? – by examining live log data gathered from the real-world public social tagging system BibSonomy. Our empirical results indicate that while some of these assumptions hold to a certain extent, other assumptions need to be reflected and viewed in a very critical light. Our observations have implications for the design of future search and other algorithms to better reflect the actual user behavior.

Categories and Subject Descriptors

H.3.4 [**Information Storage and Retrieval**]: Systems and Software—*Information Networks*

Keywords

social tagging; assumptions; social sharing; folksonomy; bookmarking; tagging; behavior

1. INTRODUCTION

Social tagging systems such as Delicious, BibSonomy or Flickr have attracted the interest of our research community for almost a decade [24, 12]. Significant advances have been made with regard to our understanding about the emergent, individual and collective processes that can be observed in such systems [20]. Useful algorithms for retrieval [16] and classification [37] have been developed that exploit the rich fabric of social tagging systems for facilitating search or navigation. Other work has focused on the extraction or stabilization of semantics [11, 7].

While this work has significantly increased our ability to describe, model and utilize social tagging systems, we can also assert that our community has built their work on certain assumptions that have gone largely unchallenged in the past, such as (i) tagging systems are social (as in *social* tagging systems) (ii) the key components of folksonomies - i.e., users, tags and resources - are equally important structural elements of social tagging systems and (iii) popularity in terms of posts means importance. Understanding whether these assumptions hold has important implications for future advances in this field, for example the FolkRank [16] algorithm builds upon assumption (ii), while the usefulness of tagging systems for facilitating knowledge transfer in groups hinges on (i). If these assumptions turned out to be wrong or applied only to very limited contexts, than FolkRank would need to be revised or the usefulness of tagging systems for knowledge transfer would have to be revisited. However, testing these assumptions on a large scale has been difficult historically, as some of these assumptions can only be studied by exploring detailed usage logs that exhibit quantifiable trails of user behavior.

In this paper, we aim to explore these assumptions in a systematic manner by investigating complete usage logs of BibSonomy[1], a social tagging system our research group has

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

been developing and running for several years. Not only due to its availability for research purposes[2], BibSonomy has been subject to several studies – e.g., on tag [18] or item recommendation [4], on network properties [7], on search and ranking [17], its connection to scientometrics [13] or the connection between pragmatic and emergent semantics [28]. Yet, all these studies have dealt with the publicly visible manifestation of the system's use until now, i.e., the posts of the system with their users, resources and tags. Information about user behavior (such as click streams) has been largely ignored. For the first time, we will investigate the actual usage of BibSonomy by analyzing server logs that contain every request that has been made to the system. The study of these logs allows us to assess which resources have been requested, the frequency of such requests, and whether users visit their own content or that of others.

Research Questions and Contributions. In this paper, we turn our attention on studying the following three assumptions in greater detail: First, we tackle the *social assumption* behind social tagging systems. Early work [34, 32, 29] has pointed out that personal information management may be one of the main reasons why users turn to social tagging systems, and social uses are only a secondary effect. In contrast to this view, there is also the hypothesis that tagging is social [36], and that users use social tagging systems for sharing information with others.

Second, we review what we call the *equality assumption* which can be seen as the hypothesis that all three classes of entities of a folksonomy – i.e., users, tags and resources – are equally important structural elements in a social tagging system. This assumption is inherent in the folksonomy model (see Section 3.1). A range of algorithms are based upon this idea (e.g., FolkRank) and thus, it is of importance to thoroughly investigate this assumption. If this assumption was invalid, it would have implications for the design of future search and other algorithms to better reflect the actual user behavior.

Finally, we want to investigate the *popularity assumption*. In tagging systems, popularity is used for example in tag clouds. The underlying concept is built upon the idea that the most frequently posted tags are also the most interesting ones and are therefore displayed larger in the cloud. We investigate whether these most frequently posted tags or resources are also those that get most of the attention from users in their retrieval requests.

To the best of our knowledge, this paper presents the first systematic and in-depth empirical study of the validity of fundamental assumptions about social tagging systems. We leverage actual log data gathered from the real-world public social tagging system BibSonomy for our experiments. While the results from our empirical investigations show that some of these assumptions hold to a certain extent, other assumptions need to be reviewed very critically.

Overall, our findings are relevant for researchers interested in advancing our knowledge about social tagging systems and their adoption as well as to system engineers who design and operate social tagging systems 'in-the-wild'.

Structure of the Paper. In Section 2, we provide a short overview of the related work. Section 3 covers the folksonomy model as well as our experimental setup including a description of the used dataset gathered from BibSonomy. In Section 4 we evaluate the three assumptions about social tagging systems. We also give a short discussion of the results for each assumption. Finally, in Section 5 we summarize our findings and point to future research.

2. RELATED WORK

In this section we will discuss related literature on the investigation of tagging systems and log file analysis in general. Related work that is specifically relevant for one of the assumptions introduced in Section 1 will be discussed in greater detail later in the corresponding context.

In 1998, early work by Abrams et al. [1] already discussed the management of website bookmarking long before the rise of social tagging on the Web. The authors analyzed several aspects of website bookmarking using data collected from a user survey and bookmark files from participants. The study shows that one motivation for creating large bookmark collections is sharing these bookmarks (still via email back then). Furthermore, the study points out that keeping the overview over all bookmarks is difficult for users and that categorizing them using folders is unstable and time consuming. The study also shows, that almost half of the users sporadically take time to organize their bookmarks, and only 26% of the respondents keep unorganized collections. With regard to the retrieval of bookmarks (the actual revisit of the corresponding website) the study shows, that on average out of a collection 67% (96%) of the bookmarks had been visited within the last 6 month (last year). Thus bookmarks are attributed an archival purpose rather than temporary importance.

The discussion around social tagging and emerging folksonomies began in late 2004, when the term *folksonomy* was first mentioned by Van der Wal [35] and continued in 2005 in various blog posts and papers [33]. The first review about social tagging systems was provided by Mathes [24]. He noted that social tagging systems allow a much greater variability in organizing content than formal classifications can provide. He also pointed out that users do not necessarily need to agree on a hierarchy of tags but only on a general sense about the meaning of tags, which is one of the main concepts of social tagging systems. Mathes was also the first to hypothesize that tag distributions may emerge to power law distributions which can characterize the semantic stabilization of such systems. Furthermore, Mathes also identified some very strong potentials of social tagging systems: serendipitous browsing and handling desire lines by having opportunities to easily adapt to user vocabulary changes.

The investigation of web logs is a method that has previously mainly been used to investigate the query behavior in search engines and the usage of digital libraries. Such analysis helps to understand a system's users and its results can help webmasters to tailor their systems more specifically to the users' needs. A survey on such works is given by Agosti et al. in [2]. In the case of search engines, aspects like how

[2] `http://www.kde.cs.uni-kassel.de/bibsonomy/dumps`

users make their requests to the search engine or how users interact with the search engine and how the search engine organizes the results have been analyzed. Examples for the analysis of library uses can be found in the works of Nicholas (among others e.g., [27]). For example, the share of visits to particular types of pages (e.g., pages with the full text of an article vs. pages containing only an abstract, etc.), viewing times of particular resources and session lengths have been examined. Tagging systems exhibit aspects of both search engines and libraries. While they are indeed collections of resources with description and categories, however not professionally organized like in a digital library, they are organized by users in their individual fashion of assigning tags and entering meta data. On the other hand, the data is clearly more structured than general data on the Web as posts are constructed according to a specified template.

Carman et al. [6] combine tagging data with log data from search engines and compare the distribution of tags to that of query terms in search. They find a large overlap in the systems' vocabularies and correlations between the frequency distributions of queries and tags to the same URLs. However, they also provide evidence that both tag and query term samples do not come from the same distribution.

The work closest to the one provided in this article is by Millen and Feinberg [25] who investigated user logs of the social tagging system "dogear" which was internally used at IBM. The authors were particularly interested in social navigation in the system. They found strong evidence that social navigation – i.e., users who are regularly looking at bookmark collections of other people – is a fundamental part of the social tagging system. They also found a positive correlation between the assignment frequency of a tag in the system and the frequency it is used for browsing. However, their work focuses on a local social tagging system located inside a company network. Therefore, it represents a private system where users only tag resources inside the company's field of interest and hence, the results are hard to judge for real world tagging system. Contrary, in this work we focus on the publicly available system BibSonomy to overcome this limitation. While we not only extend the analyses by Millen and Feinberg [25], by investigating a series of assumptions about social tagging systems, we also benefit from long-time log data allowing us to get a clearer overview over actual user behavior in an already established social tagging system.

3. EXPERIMENTAL SETUP

In this section we establish some fundamental constructs used in folksonomies and the methods we use in our experiments. Furthermore, we will describe the data that is the basis of our investigations.

3.1 Tagging and Folksonomies

The basic idea of a tagging system is that each user can post resources and annotate these resources with freely chosen keywords (tags). The structure emerging from these activities is called a *folksonomy*. It is modeled as a quadruple $\mathbb{F} := (U, T, R, Y)$, where U, T, and R are the finite sets of all users, tags, and resources, respectively, and $Y \subseteq U \times T \times R$ is a ternary relation between them, whose elements are called *tag assignments*. Hereby, $(u, t, r) \in Y$ means that the user u has annotated the resource r with the tag t (cf. [17]).

3.2 Dataset

The datasets used in this paper are generated from the social tagging and publication sharing system BibSonomy [3]. The system allows users to store and share links to websites as well as (mostly) scientific publications. BibSonomy was created in 2006 as a practical implementation of the folksonomy model [15]. BibSonomy offers the folksonomy-typical options to query for posts. A user can for example request to see all posts that are annotated with one or several tags,[3] that have been posted by a certain user,[4] or use a combination of user and tag restrictions.[5] For each resource, BibSonomy has a page, that lists its tags and users from all posts.[6] Publication posts have a details page,[7] that shows the meta data of the publication (as entered by the user who created the post) and offers export options. Posts of bookmarked websites can also contain meta data (like a description of the bookmarked website), but requests to a bookmark are usually conducted by just clicking on a post's title to reach the website. Another option to find posts is a full text search.[8] Moreover, users can form groups or declare their friendship to other users. Both friendships and groups are used in the visibility concept of posts. BibSonomy offers many further features like discussion forums or cv pages, that exceed the usual tagging system functionality. Therefore, such features have been excluded from our experiments.

Due to its high rank in search engines, BibSonomy is a popular target for spammers. Spammers are users who try to store links to their own or to advertisement sites in the system to increase the number of clicks and their rank in search engines. For detecting spammers, BibSonomy uses a learning classifier [19] as well as manual classification by the system's administrators. In all experiments, we only consider data of users that have not been flagged as spammers.

Next, we describe in detail the two different datasets we used in our study. We restricted the datasets to data that had been created in the period between the start of BibSonomy in 2006 and the end of 2011, since early in 2012 the login mechanism was modified, which introduced significant changes to our logging infrastructure.

User and Content Data. In our analysis we used typical folksonomy data: all tags assigned to both publications and bookmarks, the resources themselves, all users (non-spammers, see above) who signed up until the end of 2011 as well as all data about formed groups and friendships. In this period of time, 852 172 people registered a user account. 266 494 users were classified as non-spammers and they created 5 715 781 annotations in 551 606 bookmark posts and 2 464 856 publication posts.

[3]e.g. http://www.bibsonomy.org/tag/web+mining
[4]e.g. http://www.bibsonomy.org/user/hotho
[5]e.g. http://www.bibsonomy.org/user/hotho/web+mining
[6]e.g. http://www.bibsonomy.org/bibtex/157fe43734b18909a24bf5bf6608d2a09
[7]e.g. http://www.bibsonomy.org/bibtex/157fe43734b18909a24bf5bf6608d2a09/hotho
[8]e.g. http://www.bibsonomy.org/search/web+mining

Request Log Data. The log files include all HTTP requests by registered users, visitors and Web robots to BibSonomy from 2006 till 2011. In addition to the common request attributes like IP address, user agent, date and path we also logged the session identifier used by our web application layer as well as the content of a cookie containing the name of the logged-in user[9]. Between 2006 and 2011 we have collected over 2.5 billion requests. Before we analyzed the request data from the log files, we filtered out all requests to additional resources including Cascading Style Sheets files, JavaScript files, and images. To identify and filter out web bots, we first used the user agent attribute of each request. We collected user agents of known bots from various online databases and removed the matching requests. Since the user agent attribute in a HTTP request can be modified, further cleaning was needed to identify the remaining disguised bots. To that end, we first obtained the host names of the IP addresses and removed all requests from IP addresses of cloud computing services like Amazon Web Service. As a last cleaning step, we removed all sessions of logged-out users that contained at least four requests per second. The rationale behind this step is, that it is very unlikely that a human user can request four Web pages in a second even with tabbed browsing. After removing requests to pages that are irrelevant to our study (like help pages and administration pages), the remaining dataset contained about 170 million requests that were considered for the analysis in this paper.

The log files do not contain requests to bookmarks because the link of a bookmark points directly to the web resource and not to BibSonomy, which means that a request to that link is not logged on our server. Therefore, we must restrict some of our experiments exclusively to publication requests.

3.3 Methodology

To test the hypotheses that are underlying our assumptions about social tagging systems, we employ several statistical measures. In this section we review them before we begin the discussion of their results in the next section.

Jensen–Shannon Divergence. The Jensen-Shannon divergence $JS_b(P||Q)$ is a measure to evaluate how similar two distributions are. It is a symmetrized derivative of the (non-symmetric) Kullblack-Leibler divergence $KL_b(P||Q)$ [21] (using base $b > 0$ for the logarithm).

Given two discrete distributions P and Q, the Jensen-Shannon divergence is defined as follows:

$$JS_b(P||Q) := \tfrac{1}{2}KL_b(P||M) + \tfrac{1}{2}KL_b(Q||M),$$

where

$$M := \tfrac{1}{2}(P+Q) \quad \text{and} \quad KL_b(P||Q) := \sum_i P(i) \cdot \log_b \frac{P(i)}{Q(i)} \; .$$

The Jensen-Shannon divergence is well defined for any two distributions, as opposed to $KL_b(P||Q)$, which cannot be computed when there are events i that have probability zero in Q but not in P. The value of JS_b is always non-negative and depends on the choice of the logarithm base, however, only by a constant factor. The particular choice of the dual logarithm results in an upper bound of 1 for JS_2.

[9]The system uses this cookie to re-login the user automatically after their session has expired.

Correlation Coefficients. To measure the correlation between request and post frequencies (Section 4.3), we use Pearson's as well as Spearman's correlation coefficient and review both of them here.

Pearson's r. Given two discrete random variables X and Y with sample pairs $(x_1, y_1), \ldots, (x_n, y_n)$, Pearson's correlation coefficient $r_{X,Y}$ or just r can be calculated by

$$r_{X,Y} := \frac{\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \overline{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \overline{y})^2}} \qquad (1)$$

Pearson's r is bounded between -1 and 1, where 1 (-1) denotes a strong (inverse) linear correlation between X and Y, and $r = 0$ means that X and Y are not correlated.

Spearman's ρ. Spearman's $\rho_{X,Y}$ works similar as Pearson's correlation coefficient, but instead of taking into account the actual *values* of x_i and y_i, this correlation coefficient focuses on their *ranks* $rg(x_i)$ and $rg(y_i)$ that x_i and y_i assume in ordered lists of all values of X and of Y respectively. Spearman's ρ is thus computed as Pearson's r of these ranks. Duplicate values, which would assume the same rank position, are instead assigned an average rank. The Spearman correlation of 1 (-1) denotes a strong (inverse) *monotonic* correlation between X and Y. Other than Pearson's r, Spearman's correlation value determines how well the relationship of X and Y can be described using any monotonic function instead of only a linear function.

4. ASSUMPTIONS

In this section, we present our analysis of three major assumptions about social tagging systems along the collected data from our social bookmarking system BibSonomy. Each assumption will be introduced and discussed in depth and we will use the data of our system to check the validity of our claims.

4.1 The Social Assumption

How social is a social tagging system? The assumption, that tagging systems are indeed part of the family of social software is based on the idea, that users of such systems tag their resources publicly and can see the collections of other users and henceforth, might also use resources posted by others. The social aspect of tagging has been praised early in the history of tagging systems. Mathes [24] stated already in 2004 that folksonomies could "lower the barriers to cooperation" and in 2005, Weinberger[36] named it as one of two aspects that "make tagging highly useful." Marlow et al. [23] presented an early model for social tagging systems where they argued that social relations between users are a critical element. The authors point out that social interaction connects bookmarking activities of individuals with a rich network of shared tags, resources and users. Furthermore, Millen and Feinberg [25] found out that around 74% of all page requests in dogear – an internal social tagging service at IBM – refer to content that was bookmarked by other users. The importance of using and sharing bookmark collections in a corporate environment social tagging system were also emphasized by Damianos et al. [10].

Yet, it is not self-evident that similar observations can be made for public social tagging systems, where users use the system without direct company guidance that might influence their behavior. Contrary, users may choose to use such systems only to create their own collections and thus, ignoring the resources of other users. Vander Wal [34] already pointed out that personal information management may be one of the main reasons why people use social tagging systems and invest time for doing so which was also emphasized by Terdiman [32]. Porter [29] claimed that "Personal value precedes network value: selfish use comes before shared use". Such users can still benefit from using the system, since it allows them to collect and manage their resources (like in offline apps and managers) and to have them readily available online. Therefore, we investigate how much social interaction we can detect in the usage data of BibSonomy.

Experiments. We start by looking at all retrieval-oriented requests (this explicitly excludes edit or posting requests). We examine whether users rather request pages belonging to themselves (e.g., their own user page), pages of users they had added as friend, pages of groups they were members of, some other user's pages, or more general non-user-specific pages like tag-pages (listing all posts to one or more given tags) or pages that simply list popular resources. The number of requests to pages in each of these "ownership categories" can be found in Table 1. We can observe, that about two thirds of all requests of a logged-in user target pages in the users' own context. Users visit pages outside of their own context with about 31% of the requests to look at general pages or content of other users. This share is far below the reported 74% in [25] for a company internal tagging system. Requests to groups and friend pages both are rather infrequent (less than 2% in total). From these numbers it seems clear, that a significant part of the interaction with the system is indeed social, yet the much larger share of interaction is with the own collection.

Because tags play a key role in a tagging system, we are interested in exploring how often users query the system with tags they have already used in one of their own posts. Therefore, we considered requests to pages that are associated with at least one tag – e.g., tag pages, listing all resources tagged with one or more specified tags, or user-tag pages, listing only posts of a specified user that additionally have certain tags. We found, that only 34% (99 836 out of 289 636) of these requests contained tags that the user had not previously used to annotate a resource. Next, assuming that users retrieve their own resources using their own tags, we restricted the considered requests further by excluding requests to the users' own pages. With this restriction, the share of requests with tags outside of their own vocabulary rises to about 58% (47 921 out of 82 577 requests). We can thus observe that although users mainly query their own collections using their own tags (which seems reasonable), they more often query with other tags when they explore content of other users.

Table 1: Shares of all requests to the (logged-in) users' own content, to content from their explicit social environment (groups, that the user is a member of or friends), to content from other users or to more general pages, that do not belong to any user specifically.

	# request	% requests
user's own	1 115 754	66.91
group of user	23 100	1.39
friends of user	15 696	0.94
other users	211 315	12.67
non-user-specific	301 728	18.09

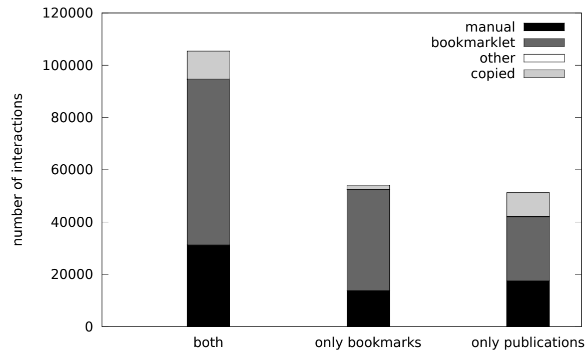
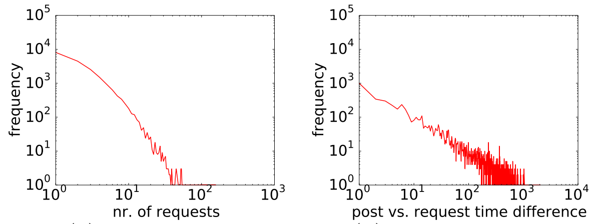


Figure 1: This histogram shows how resources are added to the system: through manual input, through a bookmarklet, through copying posts or through other methods. The three bars illustrate the ratios for both resource types together and for bookmarks and publications separately.

Next, we investigate to what extent users copy posts from others in contrast to creating new posts themselves. There are different ways to post resources to BibSonomy: (a) by manually filling in a form with all metadata of a publication or bookmark, (b) by using a bookmarklet in the browser[10] that extracts all metadata from the currently visited website or publication, (c) by copying existing entries from other users while browsing in the system, or (d) using other methods like the extraction of metadata using the ISBN of a publication.[11] For the following experiments, we had to match the posting request to store a resource into BibSonomy to the actual data of the post in the database. This task is delicate, as the only information about the posting request, that is stored in the logs is its date and user. Thus the matching must rely on these two properties only. To further cope with time asynchronicity of different machines, we only considered matches where the post's creation date in the database exactly matched the time of the request and no other resource was posted ten seconds before or after that. As Figure 1 shows, the bookmarklet feature is used most frequently (60%), followed by entering the metadata manually (30%), and only about 10% of the resources are copied from other users. Less than 1% were stored using one of the other methods. When we look at publications and bookmarks separately, we observe that the ratios between the

[10] `http://www.bibsonomy.org/buttons`

[11] Hereby we ignore options to import bookmark or publication lists from files (e.g., browser bookmark files or BibTeX) or from other resource management services. The rationale behind that is, that such imports are a means to transferring own collections to BibSonomy and for that purpose it would not be meaningful to look for resources in the collection of other users.



(a) revisit count (b) revisit time gap

Figure 2: The figures show statistics about users' revisitation behavior of their own posted publications. Both figures are visualized on a log-log scale. Figure 2(a) shows the number of times users revisit their own publications and Figure 2(b) illustrates the number of days elapsed between posting and the first request to those posts by their owners.

distinct ways of posting resources differs. Publications are copied more often than bookmarks (17% vs. 3%) and users use the bookmarklets for bookmarks (72%) more frequently than for publications (48%). One reason for these differences might be the fact, that users leave the system when they follow a bookmarked link, while they stay within BibSonomy when they check out details of a publication. Thus, using the bookmarklet is the easiest way to post a website and clicking the copy button is more likely an option for copying a visited publication. We also note, that the share of 3% of copied bookmarks is close to the 2.2% share reported in [25] for the IBM-internal system dogear, while the share for publications exceeds that value by a factor of 7.

Since a resource can only be copied if another user has already posted that resource in BibSonomy, we have to take into account whether posted resources were already in the system when a user posted them. Among the posting requests where posts were not copied from another user, only about 17% had a corresponding resource already in the database and thus could have been copied. Of these, 45% could have been found by the users using their own tags and about 40% could have been found using those tags that were later used to annotate the new post. We can conclude that, of those resources that could have been copied at the time of posting, a share of roughly 38% has indeed been copied. We consider this a relatively large share, since looking up publications in BibSonomy is only one out of many possible ways to find interesting bookmarks and publications on the web or elsewhere.

In the next experiment, we counted how often users who copied a resource used their own tags from their own vocabulary or tags of the original post to describe their new post. In 86% of all copy requests at least one tag from the own vocabulary was used. In 48% of all copies at least one of the original post's tags was adopted. In the other copy events, 44% the original posts had only special tags like "imported", that are probably not meaningful for the user copying the post.

Finally, we are interested in investigating whether users revisit their posted publications, and if so after which time they do so. Like above, we again used the restricted set of resources, and therefore only the publication posts for this experiment. The restriction to publications was necessary as requests to bookmarked (external) links do not occur in our log files. Also, we removed all posted publications by the user "dblp"[12] as this user only automatically posts resources to the system without actually requesting resources and hence, including this user would introduce a high negative bias to the values retrieved. Note, that this experiment only covers visits to the posts individually.[13] Overall, we could observe that around 40% of all posted publications are retrieved later by their owners. Particularly interesting about this observation is, that it is not in line with the work in [1] where the retrieval of browser bookmarks was investigated in a user study. There, nearly 100% of all bookmarked resources get revisited in the time span of a year. The difference might result from several factors, e.g., that using a publication again is different to revisiting a website or to the differences between online (in a tagging system) and offline (in the user's browser) storage of bookmarks. The result however suggests, that while self-retrieval is part of the system, it is not a dominant factor and the actual truth seems to lie somewhere between the observations made in [1] and those made in [25, 9, 22] who stated that the visitation of resources by other users is a major social aspect. However, we can also see in Figure 2(a) that if a publication gets revisited, it frequently only gets revisited once. This might also be a very early revisitation in order to check the post made to the system. Actually, we can see in Figure 2(b) that most of the first revisits to a page have been happening shortly after the resource has been posted and thus just might be for the purpose of checking the post.

Summary. We could clearly observe the traces of social interaction - like visiting other users content, copying posts and reusing other users' tags. In general, more requests are spent on the own collection than on other users' posts, yet the interest in other content than their own is still significant and shown in all experiments. Thus, although the explicit social network (with friends and groups) seems to play only a minor role, there is still exchange of resources and tags.

4.2 The Equality Assumption

Are users, resources, and tags equally important? A folksonomy – the structure underlying tagging systems – comprises the sets of users, resources and tags together with the tag-assignment-relation. In that model (cf. Section 3.1), users, resources, and tags are treated equally and in fact even symmetrically. The model itself has been widely accepted and many algorithms build on it, e.g., the popular FolkRank [16] or the tensor factorization method by Rendle [30]. The three classes of entities and their connections through tag assignments also build the typical folksonomy navigation idea: All entity classes are linked by the tag assignments. Thus entities can be reached by following these tag assignments, e.g., on clicking a tag one will find all posts containing that tag (and thus also all resources tagged with that tag and all users who used the tag) or on clicking a user, one will find all posts and tags of that user.

[12] `http://www.informatik.uni-trier.de/~ley/db/`

[13] Another way to make use of posted publications is to export larger parts of the collection to some reference manager or directly to BibTeX files. Since these do not specifically target one publication post, such requests do not occur in this experiment.

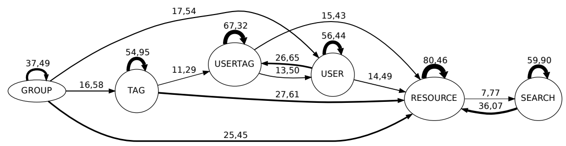


Figure 3: Transition Graph with transition probabilities relative to each node. We left out those with a transition probability less than 5%. Also, the size of the nodes is relative to the occurrence frequency of each node category. The nodes USER, RESOURCE, SEARCH, TAG, and GROUP correspond to the columns in Table 2, with the difference that we used an extra node (USERTAG) for requests to both a user and tags.

By the *equality assumption* we describe the assumption, that indeed users, resources and tags are equally important in the system. A counterargument to that assumption is the fact that tag assignments usually occur in groups, which are represented by the posts of a tagging system. A post is created by a user who assigns several tags to one resource. One post can contain several tags, but only exactly one resource and one posting user. When discussing navigation within folksonomies, resources are usually regarded as targets of queries and often tags receive the larger interest, compared to users as the navigational means to find or retrieve these resources. To check this assumption in practice, we count how many of all the requests in the system target specifically some user, tag or resource page.

Experiments. Table 2 shows the total number of retrieval requests in the first and their percentage shares in the second line; *user* subsumes all requests to post lists of a specific user, *resource* all such requests to particular resources and *tag* all requests to lists of a specific tag. Like in many social networks, users can form groups in BibSonomy and we additionally list all requests to post lists of a specific *group*. The page of a group provides an aggregated view of all posts of the group's members. Under *search* we subsume all those requests that are actual textual search queries (requests to the text search, not tag queries). Search requests can be made with arbitrary query terms and they can be restricted to users or groups. As the search function however stands out from the usual folksonomy navigation model of users, resources, and tags connected by tag assignments as links, we consider search requests limited to a specific user not as user page.

Still, some requests can belong to more than one column in Table 2. For example, a request to all posts of a user that have a specific tag would be counted both as a user request and as a tag request. Thus, the percentage shares in the second line of the table do not add up to 100%. Requests that are not specific to some entity (like the start page) are ignored. Clearly, we see that request shares are not equal at all. There are about 3.68 times more requests to specific users than to specific tags. Even tag queries and search queries added up do not make up half the number of requests to users. Moreover, requests to groups could also be considered as requests to users.

The use of a tagging system consists of managing, maintaining and retrieving posts in the own collection as well as in browsing and retrieving posts from other users. It is therefore reasonable to examine these two aspects separately. To that end, we conduct the same experiment as above, however, this time only those requests are counted that do not refer to pages of the currently logged-in user. These numbers are reported in the third and fourth line of Table 2. The last two rows count only such requests that have been made by users to their own collections. We can observe that the shares of the requests change dramatically. Among the request to pages outside the users' own collections there are still more requests to user pages, than to tags or individual resources, yet the ratio has gotten much smaller. It is especially very interesting, that the search requests now outnumber the tag requests and even the user requests. In the case of requests to the users' own pages, comparing the counts to user pages and tag pages is not meaningful, as then almost each tag query automatically is also a user query. However, we can also compare search and tag requests. We find, that in their own collections users tend to query a lot more with tags than through the (more general) search, while for browsing the whole system, search queries dominate the tag requests.

Next, we look at transition probabilities from one page category to another (see Figure 3). We calculated the local transition probabilities as well as a relative size of the nodes corresponding to their occurrence frequency as described in Table 2. We also extracted the tag requests on a user page as an extra node, because it adds heavy weight to either of them. If a transition probability accounted to less than 5%, we neglected it for the sake of readability. We clearly see that self-transitions are dominant for all categories. This suggests that users tend to same in the same type of entity in their navigational paths through BibSonomy. We also see that there are extremely many requests to user-tag pages; almost as many as there are regular user requests to user pages, and more USERTAG requests than requests on tag pages. This is likely to result from the navigational structure of BibSonomy which allows getting from a user page to a user-tag page with one click only. Thus, if one is interested in posts of a particular user, one is likely to be more interested in a tag-page that is still restricted to posts of that user, than to a more general page.

Table 2: Number of requests to pages of to different entities and to the textual search.

	user	resource	tag	group	search
# requests (any)	1 061 618	279 646	288 152	53 963	194 599
% requests (any)	65.04	17.13	17.65	3.31	11.92
# requests (not-to-self)	158 436	104 423	91 061	53 963	172 560
% requests (not-to-self)	29.80	19.64	17.13	10.15	32.45
# requests (to-self)	903 182	175 223	197 091	0	22 039
% requests (to-self)	82.06	15.92	17.91	0.00	2.00

Summary. We see that in all three cases (queries to the own collection, other requests and both together), there is *no equality* between the three entity classes of users, tags and resources. While the numbers of requests to tags and to individual resources are similar, they are dominated by the requests to user pages. This is explicitly surprising, as there are fewer user pages than tag or resource pages available in BibSonomy and still user pages get the most attention by user requests.

4.3 The Popularity Assumption

Does popularity in posts match popularity in retrieval? In tagging systems popularity is usually prominently displayed: e.g., for resources the number of posts they occur in is shown while tags are presented in tag clouds [14, 31] where frequent tags have larger fonts than less frequent ones and particularly rare tags sometimes are not displayed at all. Underlying this presentation concept is the assumption that the most frequent tags or resources are the most interesting or most important ones. Brooks and Montanez [5] pointed out that it is taken for granted that the assigned tags by users are the same as those a reader would select. Hence, the authors identified the relationship between the task of article tagging and information retrieval as an open question to investigate. To the best of our knowledge, this has not yet been investigated in a large-scale scenario, like the one which we are investigating in this section. Sinclair and Cardew-Hall investigated the usefulness of tag clouds in [31] with the help of a user study. They pointed out that using the tag cloud requires less cognitive effort than entering search queries and is useful for browsing broad categories. Tag clouds are perceived as visual summaries of the displayed resources. This indicates that the size of a tag (and thus its visibility or its probability of being noticed) is relevant for the users in their query behavior. Popularity is also reflected in other parts of tagging systems, e.g., on popular tags or resource pages where the most often used tags or the most often posted resources are displayed. Another example are tag recommender systems, that suggest to use the most popular tags in the system or the most popular tags that have previously been used for a resource. Henceforth, we want to thoroughly investigate whether the frequency of tags, resources and users in the set of all posts is reflected in the frequency of requests towards them.

To that matter we first investigate the distributions of tag frequencies in the request logs ($P^{req}_{F,Tag}$) and in the folksonomy ($P^{post}_{F,Tag}$). More precisely, $P^{req}_{F,Tag}(k)$ counts how many tags have been requested exactly k times and $P^{post}_{F,Tag}(k)$ counts how many tags have been assigned to exactly k posts. In the request distribution we do not distinguish between requests made by clicking on tags or by entering them directly into the tag search field as they are indiscernible in the logs. The latter distributions correspond to the usual node degree distributions in the folksonomy hypergraph. In previous investigations, it has been found to be fat-tailed (e.g., see [7]). We visually depict the corresponding distributions and also investigate the most applicable fit to the distributions by fitting a power law function to the data as the power law function is seen as the usual type of fat-tailed distribution for explaining tag frequency distributions in social tagging systems [7]. This presence of a power law distribution produced by tagging is often also seen as an indicator of stability as the main property of power law distributions is their scale invariance. In this work we compare the power law fits for tag distribution generated by requesting tags (i.e., tags used for the navigation) with those generated by tags in posts. For fitting the power law function to the data we use the methods proposed by Clauset et al. [8]. We also compare the power law fits to those of three candidate distributions: the exponential function, the lognormal function and the stretched exponential (Weibull) function. The exponential function is the absolute minimal candidate function to describe heavy-tailed distributions. The lognormal and Weibull function represent more sensible functions to describe heavy-tailed behavior.

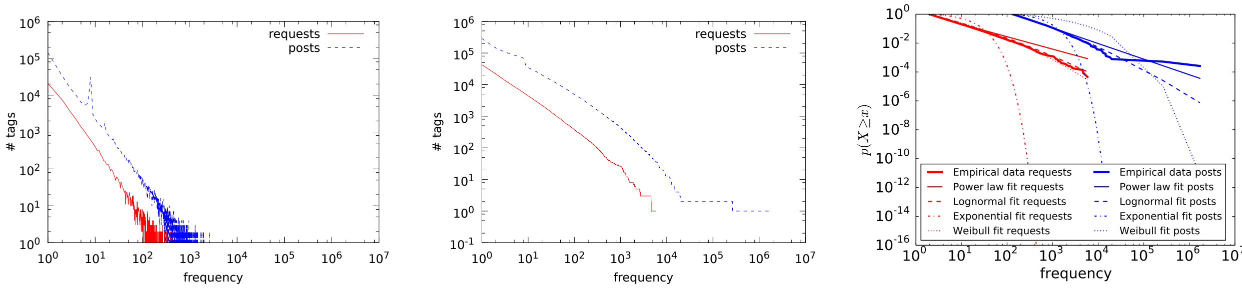
Furthermore, we also want to investigate the behavior on the level of individual tags. Hence, we compare the distributions of the tags directly (as opposed to their frequencies in the previous experiments). More precisely, we discuss the two distributions P^{req}_{Tag}, where $P^{req}_{Tag}(t)$ is the number of requests to a tag t and P^{post}_{Tag}, where $P^{post}_{Tag}(t)$ is the number of posts that the tag t occurs in.

Since BibSonomy (and other tagging system usually too) offer a regular text search next to the option to query for tags, it seems reasonable to conduct the same investigations also with queried search terms instead of queried tags. Thus, we further consider distributions $P^{req}_{F,Search}$ and P^{req}_{Search}, which are defined in the same way as $P^{req}_{F,Tag}$ and P^{req}_{Tag}, where we replace the request to a tag t by a query with a word w.

While the popularity assumption is particularly prominent for tags (through the well-known tag clouds), still the question arises, whether the distributions for the users and resources show a similar behavior. To that end, we look at analogous distributions as before with tags, i.e., P^{req}_{User} counting the requests to specific users, P^{post}_{User} counting a user's posts, P^{req}_{Pub} counting the requests to a particular publication and P^{post}_{Pub} counting the posts containing a publication. Hereby again, we restrict resources to publications (and thus omit bookmarks), as counts of visits of bookmarks are not registered in the log files (see Section 3.2).

Experiments. In these experiments we focus on distributions of tags, search terms, users and resources.

Tags. We begin our experiments by comparing the frequencies of tags used for annotating posts with the frequency of requests to tags. As mentioned above, tag frequency distributions extracted from posts in social tagging systems are known to be fat-tailed [7]. It was to be expected that the distribution of tag frequencies in the request logs has similar properties like the node degree distribution. Both distributions are shown in Figure 4 and indeed, both distributions are fat tailed and display a similar behavior. An immediate observation is, that on average, tags occur in more posts than queries. This indicates that they are used to annotate posts, but not necessarily are used for later retrieval of these posts. Surprising is the peak in the distribution $P^{post}_{F,Tag}$ at frequency 8. However, a deeper look into the folksonomy revealed, that this anomaly is in fact due to the activities of one single user, who used 28 989 tags exactly 8 times. We



(a) Distributions. (b) Cumulative Distributions. (c) Power Law Fits.

Figure 4: The frequency distributions of requests to tags ($P^{req}_{F,Tag}$) and of tags in posts of the folksonomy ($P^{post}_{F,Tag}$) in log-log scale. Displayed are the simple frequency distributions, their cumulative counterparts and fits to different standard distributions of the cumulative probability distributions.

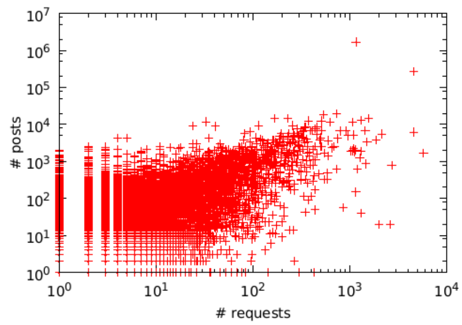


Figure 5: The scatter plot in log-log scale of the numbers of requests to a tag t vs. the number of post a tag t occurs in.

therefore ignore the peak in the further discussions.

The power law fit results in Figure 4 illustrate that the power law function is a very good fit to both $P^{post}_{F,Tag}$ ($\alpha = 2.1$ and $xmin = 133$) and $P^{req}_{F,Tag}$ ($\alpha = 1.9$ and $xmin = 2$). The visualizations and detailed log-likelihood ratio tests suggest that for $P^{post}_{F,Tag}$, the power law function is a better fit than the exponential and Weibull function. In literature, it is often stated that it is hard to distinguish the power law function from the lognormal or Weibull function [8] which is also prominent for our results. For $P^{req}_{F,Tag}$ we can also observe that the power law fit is better than the exponential fit. Interestingly, both the lognormal and the Weibull function are a better fit to the data.

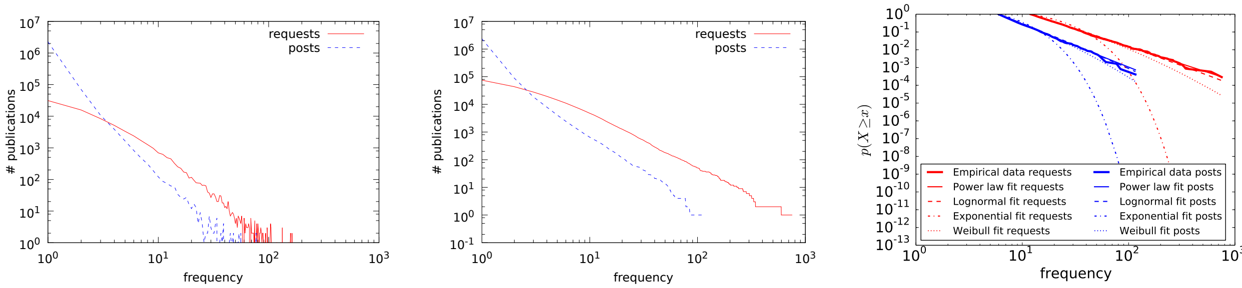
Further, we directly compare both $P^{post}_{F,Tag}$ and $P^{req}_{F,Tag}$ with each other. We computed Pearson's correlation r and Spearman's ρ between them. From the first row in Table 3 we can observe that the Pearson's and Spearman's correlation is high. An explanation for the smaller Spearman's ρ value are the fluctuations in the distributions (see Figure 4(a)) where the number of tags does no longer decrease monotonously with increasing frequency. Finally, a comparison of the distributions themselves using the Jensen-Shannon divergence confirms similarity. These results are similar to those obtained by Millen and Feinberg [25].

Figure 5 shows the scatterplot of the two individual tag distributions where each point in the diagram denotes one tag t with its number of requests $P^{req}_{Tag}(t)$ and its number of posts $P^{post}_{Tag}(t)$ as coordinates. We can see that despite the similarity in the overall behavior, on the level of individual tags there are enormous differences. Only for very frequent tags (more than 100 occurrences) one could presume a correlation between both frequency counts. To quantify the effect, the second line of Table 3 shows the correlation coefficients and the Jensen-Shannon divergence for the two distributions. In contrast to the previous distributions we can observe rather low correlation values and a much higher value for the divergence, confirming that the number of times a tag is assigned to a resource in a post and the number of times a tag is queried are only very mildly correlated. A closer look at the data revealed, that many tags which have been used in posts were never queried at all and several tags that have been queried were not assigned to any post. Therefore, we conduct the same experiment as before but we specifically ignore tags that only occur in one of the two distributions.

We yield two distributions ${}^{\emptyset}P^{req}_{Tag}$ and ${}^{\emptyset}P^{post}_{Tag}$. Hereby, we reduce the number of considered tags significantly to only 9.74%. Their distributions' correlations and divergence can

Table 3: Pearson's correlation coefficient r, Spearman's rank correlation coefficient ρ and the Jensen-Shannon divergence JS_2 for two distributions in each line. In each line a distribution P^{req}_{Entity} (where entity is either tag, search term, user or resource) of requests (or their frequencies ($P^{req}_{F,Entity}$)) is compared to one distribution P^{post}_{Entity} of posts in the folksonomy (or their frequencies ($P^{post}_{F,Entity}$)).

requests	folksonomy	r	ρ	JS_2
$P^{req}_{F,Tag}$	$P^{post}_{F,Tag}$	0.976	0.584	0.044
P^{req}_{Tag}	P^{post}_{Tag}	0.164	0.065	0.527
${}^{\emptyset}P^{req}_{Tag}$	${}^{\emptyset}P^{post}_{Tag}$	0.164	0.510	0.427
$P^{req}_{F,Search}$	$P^{post}_{F,Tag}$	0.975	0.619	0.045
P^{req}_{Search}	P^{post}_{Tag}	0.022	-0.257	0.653
${}^{\emptyset}P^{req}_{Search}$	${}^{\emptyset}P^{post}_{Tag}$	0.020	0.425	0.533
$P^{req}_{F,User}$	$P^{post}_{F,User}$	0.980	0.264	0.132
P^{req}_{User}	P^{post}_{User}	0.038	0.451	0.713
${}^{\emptyset}P^{req}_{User}$	${}^{\emptyset}P^{post}_{User}$	0.037	0.710	0.689
$P^{req}_{F,Pub}$	$P^{post}_{F,Pub}$	0.868	0.793	0.298
P^{req}_{Pub}	P^{post}_{Pub}	0.488	0.128	0.881
${}^{\emptyset}P^{req}_{Pub}$	${}^{\emptyset}P^{post}_{Pub}$	0.560	0.238	0.153



(a) Distributions. (b) Cumulative Distributions. (c) Power Law Fits.

Figure 6: The frequency distribution of requests to publications ($P^{req}_{F,Pub}$) and of publications in posts of the folksonomy ($P^{post}_{F,Pub}$) in log-log scale. Displayed are the simple frequency distributions, their cumulative counterparts and fits to different standard distributions of the cumulative probability distributions.

be found in the third line of Table 3 and we observe that the limitation to such "active" tags yields higher correlation and less divergence. However, the active tag's rankings exhibit far less ties than the full set of tags which might result in the significantly rising Spearman's ρ.

Search terms. The fourth, fifth, and sixth row of Table 3 show the correlation and divergence of the distributions of queried terms to the according distributions of tags in the folksonomy. We observe that the frequency distributions are similarly strongly correlated as the requested tag distributions. In the comparison of the counts per term (P^{req}_{Search} and P^{post}_{Tag}), we observe less, yet comparably high correlation than before. The result seems plausible as search terms can be picked more freely than tags since the text search will retrieve results when the queried words are part of a post (including its resource) and not only the tags of a post.

Users and Publications. We now focus on distributions for both users and resources. The correlation results are depicted in lines seven through twelve and the publication distributions (exemplary – the ones for users have been omitted due to space limitations) are illustrated in Figure 6. In general, we observe similar tendencies as with tags. The frequency distributions $P^{req}_{F,User}$ and $P^{post}_{F,User}$ as well as $P^{req}_{F,Pub}$ and $P^{post}_{F,Pub}$ are similar and yield relatively high correlation according to Spearman's ρ. Their Jensen-Shannon divergence is higher than it was for tags (and search terms), yet still the distributions are relatively similar. The highest divergence occurs with the frequencies of publications. Notable in both cases (users and publications) is that the frequency distributions of posts and requests are no longer "parallel" as they were in the case of tags (compare Figures 4 and 6). Since the distributions are for the most part monotonically decreasing, the rank correlation is high, yet the divergence rises in comparison to the tag distributions. We also observe that removing "inactive" entities (users and resources without posts or requests) yields higher rank correlation and lower divergence. This is most extreme in the case of publications where the elimination leaves only about 2.76% of the original set of publications. For this ratio, automatically posting accounts might be a reason. Such accounts post content for other users to find. However, they never retrieve them themselves and thus, much of the content stays unrequested. For users, the rank correlation is the strongest while divergence is high. This indicates that users with many posts indeed tend to make more requests, but not proportionally more. The power law fits for the publication distributions are for both posts ($\alpha = 3.6$, $xmin = 14$) and requests ($\alpha = 2.9$, $xmin = 12$) decent fits. However, this time we can only see that the fits are statistically significant better than those of the exponential function, while it is difficult to distinguish the fits of the lognormal and Weibull function with the power law fit.

Summary. Our obtained results clearly support the initial assumption. It seems that among actively used tags, those that are used more often in posts are also queried more frequently, although not proportionally more often. It is surprising that despite the fact that tag-clouds are displayed in BibSonomy and users can click tags to find according resources, the choice for the tag to query is not more related to its popularity in the system. The removal of tags that do not occur in both requests and posts did boost the (Spearman) correlation. For operators of a tagging system, this would indicate that it is reasonable to exclude rarely requested tags completely from tag clouds. The power law results indicate that there may be difference in the underlying process of assigning and requesting tags even though both distributions seem to be heavy-tailed. However, no clear statement is possible as the power law, lognormal and Weibull functions are all reasoned on similar behavior. Starting with the observations made in [26], we want to thoroughly investigate the underlying process in future work.

Overall, we find that in general, the frequency distributions of the entities in the tagging systems (also including search terms, users and publications) are relatively similar and they seem to be heavy-tailed. Entities that are requested very often also tend to occur in more posts. However, there are many items that are rarely requested and still occur frequently in posts.

5. CONCLUSIONS

In this work we have analyzed and discussed three typical assumptions about social tagging systems by leveraging log data from the social tagging system BibSonomy to evaluate their validity. With regard to the *social assumption* – postulating that social tagging systems are indeed social by the existence of social sharing – we found that this assumption is valid as a decent amount of exchange and visiting of resources and tags can be observed. Second, we found that the *equality assumption*, which states that users, resources and tags are equally important in a folksonomy, clearly does not hold as users play the most important part in the information retrieval processes even though there are fewer users than tags or resources in the folksonomy itself. Finally, we could see that the *popularity assumption* – stating that popular entities are the most important ones for retrieval – only holds true for the more actively used entities and the assumption as such is too strong as the popularity of entities in posts is not well reflected in their frequency of being requested.

As a consequence, some of these assumptions may need rethinking and also, algorithms building on top of such hypotheses have to be reviewed. For example, FolkRank [16] builds upon the equality assumption which we found not to hold at all. Further, tag clouds might profit from the exclusion of rarely requested tags in contrast to rarely used tags as our investigations on the *popularity assumption* showed. The design of tagging systems might also need to focus more on the navigation of user pages.

In future investigations we will further address the problem of automatic imports by artificial users. As in a folksonomy, users are connected to their resources and tags and since the generated content of artificial users is visited, retrieved and copied by human users, the removal of particular users is a non-trivial problem. Another aspect for deeper investigatation is the use of resources. Often, publications are stored in BibSonomy to cite them later. It is common to select some tag and to export all own posts with that tag to some citation format (like BibTeX or Endnote). These requests are not considered so far as they are not requests to individual publications. Furthermore, we will analyze individual user behavior. In this paper we have investigated structural phenomena that result from the collaborative actions of all users. We will therefore examine different types of user behavior like personal ratios of contribution and retrieval, preferences to tag querying or text search.

6. REFERENCES

[1] D. Abrams, R. Baecker, and M. Chignell. Information archiving with bookmarks: personal web space construction and organization. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '98, pages 41–48, New York, NY, USA, 1998. ACM Press/Addison-Wesley Publishing Co.

[2] M. Agosti, F. Crivellari, and G. Di Nunzio. Web log analysis: a review of a decade of studies about information acquisition, inspection and interpretation of user interaction. *Data Mining and Knowledge Discovery*, 24(3):663–696, 2012.

[3] D. Benz, A. Hotho, R. Jäschke, B. Krause, F. Mitzlaff, C. Schmitz, and G. Stumme. The social bookmark and publication management system bibsonomy. *The VLDB Journal*, 19(6):849–875, Dec. 2010.

[4] T. Bogers. *Recommender Systems for Social Bookmarking*. PhD thesis, Tilburg University, Tilburg, The Netherlands, Dec. 2009.

[5] C. H. Brooks and N. Montanez. Improved annotation of the blogosphere via autotagging and hierarchical clustering. In *Proceedings of the 15th international conference on World Wide Web*, WWW '06, pages 625–632, New York, NY, USA, 2006. ACM.

[6] M. J. Carman, M. Baillie, R. Gwadera, and F. Crestani. A statistical comparison of tag and query logs. In *Proceedings of the 32nd international ACM SIGIR conference on Research and development in information retrieval*, SIGIR '09, pages 123–130, New York, NY, USA, 2009. ACM.

[7] C. Cattuto, C. Schmitz, A. Baldassarri, V. D. P. Servedio, V. Loreto, A. Hotho, M. Grahl, and G. Stumme. Network properties of folksonomies. *AI Communications Journal, Special Issue on "Network Analysis in Natural Sciences and Engineering"*, 20(4):245–262, 2007.

[8] A. Clauset, C. R. Shalizi, and M. E. J. Newman. Power-law distributions in empirical data. *SIAM Rev.*, 51(4):661–703, Nov. 2009.

[9] L. Damianos, J. Griffith, D. Cuomo, D. Hirst, and J. Smallwood. Onomi: Social bookmarking on a corporate intranet. In *Collaborative Web Tagging Workshop at WWW2006, Edinburgh, Scotland*, May 2006.

[10] L. E. Damianos, D. Cuomo, J. Griffith, D. M. Hirst, and J. Smallwood. Exploring the adoption, utility, and social influences of social bookmarking in a corporate environment. In *Proceedings of the 40th Annual Hawaii International Conference on System Sciences*, HICSS '07, pages 86–, Washington, DC, USA, 2007. IEEE Computer Society.

[11] S. Golder and B. A. Huberman. Usage patterns of collaborative tagging systems. *Journal of Information Science*, 32(2):198–208, April 2006.

[12] S. A. Golder and B. A. Huberman. Usage patterns of collaborative tagging systems. *Journal of information science*, 32(2):198–208, 2006.

[13] S. Haustein and T. Siebenlist. Applying social bookmarking data to evaluate journal usage. *Journal of Informetrics*, 5(3):446 – 457, 2011.

[14] M. Hearst and D. Rosner. Tag clouds: Data analysis tool or social signaller? In *Hawaii International Conference on System Sciences, Proceedings of the 41st Annual*, pages 160–160, 2008.

[15] A. Hotho, R. Jäschke, C. Schmitz, and G. Stumme. BibSonomy: A social bookmark and publication sharing system. In A. de Moor, S. Polovina, and H. Delugach, editors, *Proceedings of the Conceptual Structures Tool Interoperability Workshop at the 14th International Conference on Conceptual Structures*, Aalborg, Denmark, July 2006. Aalborg University Press.

[16] A. Hotho, R. Jäschke, C. Schmitz, and G. Stumme. FolkRank: A ranking algorithm for folksonomies. In *Proc. FGIR 2006*, pages 111–114, 2006.

[17] A. Hotho, R. Jäschke, C. Schmitz, and G. Stumme. Information retrieval in folksonomies: Search and ranking. In *Proceedings of the 3rd European Semantic Web Conference*, volume 4011 of *LNCS*, pages 411–426, Budva, Montenegro, June 2006. Springer.

[18] R. Jäschke, L. Marinho, A. Hotho, L. Schmidt-Thieme, and G. Stumme. Tag recommendations in social bookmarking systems. *AI Communications*, 21(4):231–247, 2008.

[19] B. Krause, C. Schmitz, A. Hotho, and G. Stumme.

The anti-social tagger - detecting spam in social bookmarking systems. In *Proc. of the Fourth International Workshop on Adversarial Information Retrieval on the Web*, 2008.

[20] C. Körner, D. Benz, M. Strohmaier, A. Hotho, and G. Stumme. Stop thinking, start tagging - tag semantics emerge from collaborative verbosity. In *Proceedings of the 19th International World Wide Web Conference (WWW 2010)*, Raleigh, NC, USA, Apr. 2010. ACM.

[21] S. Kullback and R. A. Leibler. On information and sufficiency. *The Annals of Mathematical Statistics*, 22(1):79–86, 1951.

[22] K. Lerman and L. Jones. Social browsing on flickr. In *Proceedings of 1st International Conference on Weblogs and Social Media (ICWSM-07)*, 2007.

[23] C. Marlow, M. Naaman, D. Boyd, and M. Davis. Ht06, tagging paper, taxonomy, flickr, academic article, to read. In *Proceedings of the seventeenth conference on Hypertext and hypermedia*, HYPERTEXT '06, pages 31–40, New York, NY, USA, 2006. ACM.

[24] A. Mathes. Folksonomies: Cooperative classification and communication through shared metadata. `http://www.adammathes.com/academic/computer-mediated-communication/folksonomies.html`, June 2004. Accessed: 2013-07-11.

[25] D. R. Millen and J. Feinberg. Using social tagging to improve social navigation. In *Workshop on the Social Navigation and Community based Adaptation Technologies*, 2006.

[26] M. Mitzenmacher. A brief history of generative models for power law and lognormal distributions. *INTERNET MATHEMATICS*, 1:226–251, 2003.

[27] D. Nicholas, P. Huntington, and A. Watkinson. Scholarly journal usage: the results of deep log analysis. *Journal of Documentation*, 61(2):248–280, 2005.

[28] T. Niebler, P. Singer, D. Benz, C. Körner, M. Strohmaier, and A. Hotho. How tagging pragmatics influence tag sense discovery in social annotation systems. In *Proceedings of the 35th European Conference on Advances in Information Retrieval*, ECIR'13, pages 86–97, Berlin, Heidelberg, 2013. Springer-Verlag.

[29] J. Porter. Learning More about Structured Blogging. `http://bokardo.com/archives/learning-more-about-structured-blogging/`, 2005. Accessed: 2013-08-12.

[30] S. Rendle, L. Balby Marinho, A. Nanopoulos, and L. Schmidt-Thieme. Learning optimal ranking with tensor factorization for tag recommendation. In *Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 727–736. ACM, 2009.

[31] J. Sinclair and M. Cardew-Hall. The folksonomy tag cloud: when is it useful? *Journal of Information Science*, 34(1):15–29, 2008.

[32] D. Terdiman. Folksonomies tap people power. `http://www.wired.com/science/discoveries/news/2005/02/66456?currentPage=all`, Jan. 2005. Accessed: 2013-08-12.

[33] J. Trant. Studying social tagging and folksonomy: A review and framework. *J. Digit. Inf.*, 10(1), 2009.

[34] T. Vander Wal. Tagging for fun and finding. `http://okcancel.com/archives/article/2005/07/tagging-for-fun-and-finding.html`, July 2005. Accessed: 2013-08-12.

[35] T. V. Wal. Folksonomy, 2007. http://vanderwal.net/folksonomy.html.

[36] D. Weinberger. Tagging and why it matters. *SSRN eLibrary*, 2005.

[37] A. Zubiaga, C. Körner, and M. Strohmaier. Tags vs shelves: from social tagging to social classification. In *Proceedings of the 22nd ACM conference on Hypertext and hypermedia*, HT '11, pages 93–102, New York, NY, USA, 2011. ACM.