

Evaluating Assumptions about Social Tagging ^{*}

A Study of User Behavior in BibSonomy

Stephan Doerfel¹, Daniel Zoller², Philipp Singer³, Thomas Niebler²,
Andreas Hotho², and Markus Strohmaier^{3,4}

¹ ITeG & Knowledge and Data Engineering Group, University of Kassel (Germany)
`doerfel@cs.uni-kassel.de`

² Data Mining and Information Retrieval Group, University of Würzburg (Germany)
`{zoller, niebler, hotho}@informatik.uni-wuerzburg.de`

³ GESIS (Germany) `{philipp.singer, markus.strohmaier}@gesis.org`

⁴ University of Koblenz (Germany)

Abstract. Social tagging systems have established themselves as an important part in today’s web and have attracted the interest of our research community in a variety of investigations. Henceforth, several assumptions about social tagging systems have emerged on which our community also builds their work. Yet, testing such assumptions has been difficult due to the absence of suitable usage data in the past. In this work, we investigate and evaluate four assumptions about tagging systems by examining live server log data gathered from the 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 in a very critical light.

1 Introduction

Social tagging systems such as BibSonomy, Delicious or Flickr have attracted the interest of our research community for almost a decade. While previous research has significantly expanded our expertise to describe [4] and model [2], social tagging systems, the community has also built their work on certain assumptions about usage patterns in these systems, which have emerged over time. For such assumptions, arguments and evidence have been discussed, though it is not clear to which degree they remain valid in actual tagging systems. Only a few studies have analyzed user behavior in social tagging systems to better understand such assumptions, either by (i) conducting user surveys (e.g., [5]) or by (ii) tapping into the rich corpus of tagging data (i.e., the posts) that is available on the web (e.g., [2]). However, such studies lack of detailed data how users actually

^{*} Extended Abstract for Work-in-Progress.

Copyright © 2014 by the paper’s authors. *Copying permitted only for private and academic purposes.* In: T. Seidl, M. Hassani, C. Beecks (Eds.): Proceedings of the LWA 2014 Workshops: KDML, IR, FGWM, Aachen, Germany, 8-10 September 2014, published at <http://ceur-ws.org>

request information. In this paper we overcome these drawbacks by presenting and thoroughly investigating a detailed usage log of the real-world, open social tagging system BibSonomy.⁵

2 Assumptions and Results

The Social Assumption. Assuming that social tagging systems are social, we measure to which degree users collaboratively share resources and we discuss evidence for the interest of users in the content of others. Details of this analysis can be found in [3].

The Retrieval Assumption. For the retrieval assumption we investigate whether users store resources in BibSonomy for later retrieval. We discover that while users post a large number of resources and tags to BibSonomy, they only retrieve a rather small fraction of them later.

The Equality Assumption. The equality assumption claims that the three sets of entities in a tagging system – users, tags, and resources – are equally important for navigation and retrieval. However, we find a strong *inequality* in the use of these entity sets: in BibSonomy, requests to user pages dominate the number of requests to tags and to resources.

The Popularity Assumption. Finally, we test whether the popularity of users, tags, and resources in posts is matched by their popularity in retrieval. We observe common usage patterns in posting and requesting behavior on an aggregate level. The patterns are less pronounced on an individual level.

Acknowledgments. This work is in part funded by the DFG through the PoSTs II project.

References

1. Benz, D., Hotho, A., Jäschke, R., Krause, B., Mitzlaff, F., Schmitz, C., Stumme, G.: The social bookmark and publication management system BibSonomy. *The VLDB Journal* 19(6), 849–875 (Dec 2010)
2. Cattuto, C., Schmitz, C., Baldassarri, A., Servedio, V.D.P., Loreto, V., Hotho, A., Grahl, M., Stumme, G.: Network properties of folksonomies. *AI Communications Journal*, Special Issue on “Network Analysis in Natural Sciences and Engineering” 20(4), 245–262 (2007)
3. Doerfel, S., Zoller, D., Singer, P., Niebler, T., Strohmaier, M., Hotho, A.: How social is social tagging? In: *Proceedings of the 23rd International World Wide Web Conference. WWW 2014*, ACM, New York, NY, USA (2014)
4. Golder, S.A., Huberman, B.A.: Usage patterns of collaborative tagging systems. *Journal of information science* 32(2), 198–208 (April 2006)
5. Heckner, M., Heilemann, M., Wolff, C.: Personal information management vs. resource sharing: Towards a model of information behaviour in social tagging systems. In: *Proceedings of the 3rd International Conference on Weblogs and Social Media. ICWSM '09*, San Jose, CA, USA (May 2009)

⁵ <http://www.bibsonomy.org/>, see [1] for a detailed description and various analyses.