Journal Evaluation and Science 2.0: Using Social Bookmarks to Analyze Reader Perception

Stefanie Haustein*[1,2], Evgeni Golov[2], Kathleen Luckanus[2], Sabrina Reher[2], Jens Terliesner[2]

[1] *s.haustein@fz-juelich.de*
Forschungszentrum Jülich, Central Library
52425 Jülich (Germany)

[2] *{evgeni.golov, kathleen.luckanus, sabrina.reher, jens.terliesner}@uni-duesseldorf.de*
Heinrich-Heine-University Düsseldorf, Department of Information Science
Universitätsstraße 1, 40225 Düsseldorf (Germany)

Introduction

Web 2.0 technologies are blazing a trail into science: specialized social bookmarking services allow researchers to store and share literature online [SHNEIDERMAN, 2008]. In science 2.0, academics generate information about resources, i.e. description and usage. Since publishers deny access to global download statistics, we propose the application of social bookmarking data to journal evaluation: the number of bookmarks could function as an alternative indicator of journal usage. Tags assigned to the articles reveal the readers' perspectives on the journal content.

Method

A set of 45 solid state physics journals was defined. The bibliographic data for the 168.109 documents published in these periodicals from 2004 to 2008 was downloaded from the Science Citation Index (SCI). To retrieve bookmarking data, we chose CiteULike.org and BibSonomy.org as data sources. These social bookmarking systems support the sharing of bibliographic references to scientific publications, and, as such, target academic users [HAMMOND et al., 2005].
Since a great share of the metadata of the bookmarks was incomplete or erroneous, we defined three different search strategies to obtain all the bookmarks assigned to our journal set: searching for the titles of the periodicals plus their common abbreviations, ISSNs, and DOIs of the articles. This information was collected for all of the 45 periodicals via SCI, http://zdb-opac.de, Ulrich's Periodicals Directory, http://dx.doi.org and http://crossref.org. In BibSonomy, bookmarks were retrieved in XML-format with a fulltext search via the API. As CiteULike does not offer an API, results were retrieved with field searches (for title, ISSN and DOI) in RIS-format via web search.

Table 1. Indicators for the ten journals with the highest usage ratio

	Publications 2004-2008	5-Year Impact Factor	Number of Bookmarks	Usage ratio	Usage diffusion	Usage intensity
REV MOD PHYS	173	40.395	146	35.26%	98	2.4
PHYS REP	341	16.368	126	19.35%	92	1.9
REP PROG PHYS	220	12.480	62	19.09%	45	1.5
J RHEOL	347	3.008	35	9.51%	5	1.1
ACT CRYST A	326	2.098	30	7.98%	17	1.2
SOFT MATTER	654	4.890	50	6.42%	33	1.2
J STAT MECH	958	2.742	70	5.01%	53	1.5
PHYS REV E	12117	2.566	807	4.93%	291	1.4

EUR PHYS J E	707	2.306	40	4.81%	23	1.2
ANN PHYS	296	1.250	16	4.73%	12	1.1

Results

Checking for duplicates among the user names of our two result sets confirmed the assumption that users bookmark in one service only. Thus, both sets were combined. To analyze the coverage of the set of articles, we matched 3,953 bookmarks to 3,202 articles to our corrected and complemented SCI dataset by their DOIs. We propose the percentage of articles bookmarked as an indication of the *usage ratio* per journal, the number of users per journal as a measure of the *diffusion* and the average number of bookmarks per article as an indicator for the *intensity* of usage of the journal content. Table 1 shows the results for the ten journals with the highest *usage ratio*. The *5-year impact factor* and *usage ratio* correlate very strongly (r=0.937).

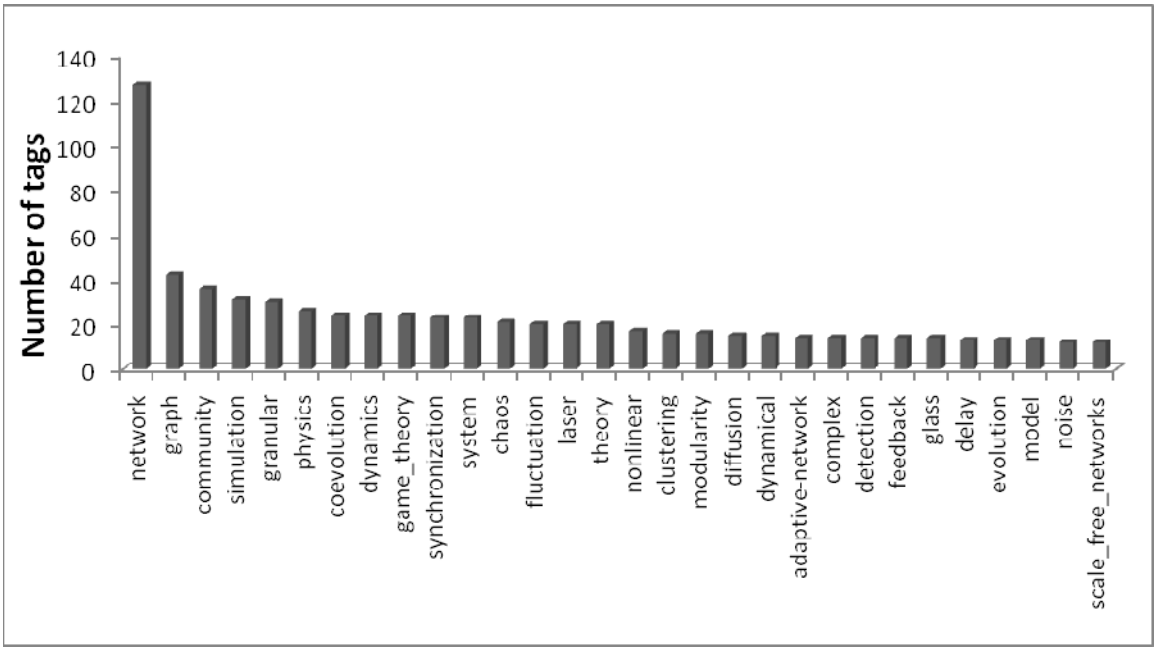

Figure 1. Most used (>11) tags assigned to *Physical Review E*.

The readers' perspectives on journal content can be described by the tags assigned to the articles. For *Physical Review E*, the journal with the highest number of bookmarks in the study, tags were adjusted in terms of merging singular and plural forms, unifying special characters and deleting content-unrelated tags. This resulted in a power-law distribution of 991 tags which were assigned 2.470 times. The first 30 can be seen in figure 1. They reflect several of the focus areas of the periodical: computational physics, granular materials and chaos theory. Thus, tags may serve as a real-time indicator of hot topics published in journals.

Outlook

Scientific social bookmarking is still in its infancy and there are different services competing for user. Bookmarking data needs improvement. Trends show a clear increase in the number of participants [BERNIUS et al., 2009]. Once a critical mass is reached, usage data will gain more significance and make it possible to indicate global reader perception with bookmarking data. The proposed measurements can then be used as supplementary indicators in multidimensional journal evaluation [JUCHEM et al., 2006; ROUSSEAU, 2002].

References

Bernius, S., Hanauske, M. and Dugall, B. (2009). Von traditioneller wissenschaftlicher Kommunikation zu "Science 2.0". *ABI-Technik* 29, 214-226.

Hammond, T., Hannay, T., Lund, B. and Scott, J. (2005). Social Bookmarking Tools (I). *D-Lib Magazine* 11.

Juchem, K., Schloegl, C. and Stock, W.G. (2006). Dimensionen der Zeitschriftenszientometrie am Beispiel von "Buch und Bibliothek". *Information Wissenschaft & Praxis* 57, 31-37.

Rousseau, R. (2002). Journal Evaluation: Technical and Practical Issues. *Library Trends* 50, 418-439.

Shneiderman B. (2008). Computer Science - Science 2.0. *Science* 319, 1349-1350.