

## Chapter 7.8

# On the Effectiveness of Social Tagging for Resource Discovery

**Dion Hoe-Lian Goh**

*Nanyang Technological University, Singapore*

**Khasfariyati Razikin**

*Nanyang Technological University, Singapore*

**Alton Y. K. Chua**

*Nanyang Technological University, Singapore*

**Chei Sian Lee**

*Nanyang Technological University, Singapore*

**Schubert Foo**

*Nanyang Technological University, Singapore*

**Jin-Cheon Na**

*Nanyang Technological University, Singapore*

**Yin-Leng Theng**

*Nanyang Technological University, Singapore*

### ABSTRACT

Social tagging is the process of assigning and sharing among users freely selected terms of resources. This approach enables users to annotate/describe resources, and also allows users to locate new resources through the collective intelligence of other users. Social tagging offers a new avenue for resource discovery as compared to taxonomies and subject directories created by experts. This chapter investigates the effectiveness of tags as

resource descriptors and is achieved using text categorization via support vector machines (SVM). Two text categorization experiments were done for this research, and tags and Web pages from del.icio.us were used. The first study concentrated on the use of terms as its features while the second used both terms and its tags as part of its feature set. The experiments yielded a macroaveraged precision, recall, and F-measure scores of 52.66%, 54.86%, and 52.05%, respectively. In terms of microaveraged values, the experiments obtained 64.76% for precision, 54.40% for recall, and 59.14% for F-measure.

DOI: 10.4018/978-1-59904-879-6.ch025

The results suggest that the tags were not always reliable indicators of the resource contents. At the same time, the results from the terms-only experiment were better compared to the experiment with both terms and tags. Implications of our work and opportunities for future work are also discussed.

## INTRODUCTION

The increasing popularity of Web 2.0-based applications has empowered users to create, publish, and share resources on the Web. Such user-generated content may include text (e.g., blogs, wikis), multimedia (e.g., YouTube), and even organization/navigational structures providing personalized access to Web content. The latter includes social bookmarking/tagging systems such as del.icio.us and Connotea.

Social tagging systems allow users to annotate links to useful Web resources by assigning keywords (tags) and possibly other metadata, facilitating their future access (Macgregor & McCulloch, 2006). These tags may further be shared by other users of the social tagging system, in effect creating a community where users can create and share tags pointing to useful Web resources. Put differently, tags function both as content organizers and discoverers. Users create and assign tags to a useful resource they come across so that it would be easy for them to retrieve that resource at a later date. At the same time, other users can use one or more of these tags created to find the resource. The same tags may also be used to discover other related and relevant resources. In addition, through tags, a user can potentially locate like-minded users who hold interests in similarly-themed resources, leading to the creation of social networks (Marlow, Naaman, Boyd, & Davis, 2006).

Social tagging provides an alternative means of organizing resources when compared with conventional methods of categorization based on

taxonomies, controlled vocabularies, faceted classification, and ontologies. Conventional methods require experts with domain knowledge and this often translates to a high cost of implementing such systems. They are also bound strictly by rules to ensure their classification schemes remain consistent (Morville, 2005). As the system becomes larger, the rules tend to be more complicated, leading to possible maintenance and accessibility issues. In contrast, the classification scheme in social tagging systems is deregulated. Instead of relying on (a few) experts, they are supported by a (possibly large) community of users. At the same time, tags are “flat,” lacking a predefined taxonomic structure, and their use relies on shared, emergent social structures and behaviors, as well as a common conceptual and linguistic understanding within the community (Marlow et al., 2006). Tags are therefore also known as “folksonomies,” short for “folk taxonomies,” suggesting that they are created by lay users, as opposed to domain experts or information professionals such as librarians, and may in fact be more effective in describing the resource

While social tagging systems have become popular, it is not known if tags created by ordinary users (as opposed to experts) are useful for the discovery of information. A few studies have investigated the use of tags as resource descriptors. Examples include comparing the use of tags against author-assigned index terms in academic papers (Kipp, 2006; Lin, Beaudoin, Bui, & Desai, 2006), examining the ability of tags to classify blogs using text categorization methods (Sun, Suryanto, & Liu, 2007), and investigating the ability of del.icio.us tags to classify Web resources in a small scale study (Razikin, Goh, Cheong, & Ow, 2007). However, to the best of our knowledge, no large scale work has been conducted with del.icio.us, one of the earliest and more popular social tagging sites. The site has a diverse set of tags and Web resources, and its main function is to store, organize, and share bookmarks among a community of users.

8 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:

[www.igi-global.com/chapter/effectiveness-social-tagging-resource-discovery/39822](http://www.igi-global.com/chapter/effectiveness-social-tagging-resource-discovery/39822)

## Related Content

---

### A Neuro-Fuzzy Partner Selection System for Business Social Networks

T. T. Wong and Loretta K.W. Sze (2012). *Handbook of Research on Business Social Networking: Organizational, Managerial, and Technological Dimensions* (pp. 98-118).

[www.irma-international.org/chapter/neuro-fuzzy-partner-selection-system/60307](http://www.irma-international.org/chapter/neuro-fuzzy-partner-selection-system/60307)

### Internet, Social Media, and Mobile Device Addiction Effects on a Workplace

Deborah S. Carstens, Jeffrey R. Miller, John A. Mahlman and Matthew J. Shaffer (2021). *International Journal of Social Media and Online Communities* (pp. 37-50).

[www.irma-international.org/article/internet-social-media-and-mobile-device-addiction-effects-on-a-workplace/280553](http://www.irma-international.org/article/internet-social-media-and-mobile-device-addiction-effects-on-a-workplace/280553)

### Social Media Tools for Quality Business Information

Sérgio Maravilhas (2018). *Social Media Marketing: Breakthroughs in Research and Practice* (pp. 883-912).

[www.irma-international.org/chapter/social-media-tools-for-quality-business-information/203335](http://www.irma-international.org/chapter/social-media-tools-for-quality-business-information/203335)

### Reasons for Adolescents' Social Media Use: Relations With Mental Health and Self-Perception

Christopher T. Barry, Jacob Matthew Briggs, Shanelle M. Briggs and Chloe L. Sidoti (2022). *International Journal of Social Media and Online Communities* (pp. 1-16).

[www.irma-international.org/article/reasons-for-adolescents-social-media-use/312180](http://www.irma-international.org/article/reasons-for-adolescents-social-media-use/312180)

### Tweets of a Feather Flock Together: An Analysis of the Impact of Twitter in a Class of Translation

Elena Alcalde Peñalver and Alexandra Santamaría Urbieta (2023). *Research Anthology on Applying Social Networking Strategies to Classrooms and Libraries* (pp. 1321-1336).

[www.irma-international.org/chapter/tweets-of-a-feather-flock-together/312982](http://www.irma-international.org/chapter/tweets-of-a-feather-flock-together/312982)