Chapter 14
Folksonomy
Creating Metadata through Collaborative Tagging

Stefan Bitzer
Georg-August-Universität Göttingen, Germany

Lars Thoroe
Georg-August-Universität Göttingen, Germany

Matthias Schumann
Georg-August-Universität Göttingen, Germany

ABSTRACT
Modern Web 2.0 technologies facilitate the collaboration and sharing of information among users, thereby enabling cooperative processes of information search. One kind of user participation is collaborative tagging, where individuals assign keywords to resources and objects on the Internet. Through the allocation of keywords, objects are enhanced with user-created metadata which results in the so-called folksonomies. This chapter focuses on the classification of tags based on function and user motivation, examines advantages and disadvantages of folksonomies, and provides a review of current applications using collaborative tagging. Future trends and potential developments are identified as they relate to the implementation of collaborative tagging in corporate settings.

INTRODUCTION
Since the introduction of social software and Web 2.0 technologies, discussion about the influence and importance of a broad range of new tools for various fields has followed (Safran, Gütl, & Helic, 2007; Boulos & Wheeler, 2007). Considerable potential for Web 2.0-based applications is seen in supporting information retrieval and classification of online resources. The Internet offers a huge amount of information varying in content, format, and quality. This diversity poses a challenge for users seeking specific information. Modern Web 2.0 technologies facilitate the collaboration and sharing of information among users, thus enabling cooperative processes of information search. While in Web 1.0 the user was not integrated into the organization of content, in today’s Del.icio.us-type folksonomy-based systems the user is actively engaged via the newly emerged tag recommenda-
Folksonomy

tion mechanisms. Such collaborative enrichment of web content enhanced with metadata is viewed as a step towards the Semantic Web (Xu, Fu, Mao, & Su, 2006).

BACKGROUND
Folksonomies and Tagging

Vander Wal (2005) defines folksonomy as the outcome of individual free tagging of online content and resources in a social environment for one’s own retrieval. With this term, Vander Wal refers to the result of a process of collaboratively assigning keywords to resources or items on the Internet, the so-called collaborative tagging. Therefore, folksonomy is often used synonymously with the terms social classification, social indexing, or social tagging (Voß, 2007). Folksonomy is a portmanteau of the words folk and taxonomy (Bateman, Brooks, & McCalla, 2006). The naming, however, is disputed. Some see it as a misnomer because of the reference to taxonomy. A classification scheme like taxonomy is strictly hierarchic and contains relations, unlike a folksonomy, which consists of a flat namespace (Mathes, 2004). The vocabulary is not preassigned, instead the users describe the information and items within their own comprehension. The purpose of folksonomies is not categorization but connecting items and expressing their meaning through personal understanding (Vander Wal, 2005).

In the context of folksonomies, three elements have to be considered (Marlow, Naaman, Boyd, & Davis, 2006), namely, resources, tags used for describing the resources, and users who assign the tags. In broad folksonomies (as in Figure 1), many users describe the same item with a term from their personal vocabulary. Hence, similar or different tags can be assigned to an object (from 1 to 5). On the basis of all assigned tags, users are able to retrieve the described object. A common example of an application of broad folksonomies is a popular social bookmarking service, Del.icio.

By contrast, in narrow folksonomies (Figure 1) there are only a few tags, mostly provided by the content creator and a group of a few people. Due to this, the number of tags and tagging persons is significantly lower than in broad folksonomies. Every tag is generally created and recorded once only, either by the content creator or a small group of selected users. Only new tags can be attributed to an object, which inhibits the possibility of counting tag frequencies. Accordingly, all tags are ranked equally and a tag distribution cannot be created. However, it can be shown via which tag users found the resource. The approach of the narrow folksonomy resembles professional indexing with controlled terms for thesauri or ontologies; in contrast, folksonomies have uncontrolled terms. Popular examples for narrow folksonomies (Cattuto, Loreto, & Pietronero, 2007) include services such as Flickr (photographs) or Technorati (blog posts).

Today folksonomies are implemented in various fields. In addition to a high diffusion of Web 2.0-based services, folksonomies are employed in corporate applications (Fichter, 2006). Capabilities have been found in indexing corporate blogs, podcasts and vodcasts (Peters, 2006); corporate social bookmarking (Damianos, Griffith, & Cuomo, 2006; Millen, Feinberg, & Kerr, 2006); and message boards (Murison, 2005).