


Tatyana Dumova
Montclair State University, USA

Richard Fiordo
University of North Dakota, USA

Volume II
Chapter 45

Critical Success Factors in the Development of Folksonomy-Based Knowledge Management Tools

Kenneth Owen
Lakehead University, Canada

Robert Willis
Vancouver Island University, Canada

ABSTRACT

This chapter examines three important aspects of folksonomies: common design factors found in folksonomies, developmental patterns of mature folksonomies, and the identification of knowledge consumer behaviors that can act as metrics for the evaluation of a small-scale folksonomy. In identifying desirable design elements, a comparative examination of tags and objects was made using a study conducted at Lakehead University. From this project, an exemplar of an effective folksonomical data structure was derived. User behavior was examined and categorized to identify behaviors that can be monitored and measured as indicators of user satisfaction. The authors analyze the structures of a folksonomy and synthesize a practical model of an effective folksonomy in the context of knowledge management.

INTRODUCTION

In today’s knowledge economy, companies struggle to find ways to collect, retain and reuse information as efficiently as possible. Control structures found in traditional Knowledge Management (KM) systems are difficult to maintain and require specialized knowledge and training to be effective (Davis, Studer, Sure, & Warren, 2005). Traditionally, knowledge management experts are hired to develop complex hierarchies and ontologies to design systems based on pre-defined information structures (such as categories and relationships). Maintaining these systems and ensuring they continue to match the needs of an organization is a skilled art. If managed poorly, these systems have the potential to miscategorize and lose data. Additionally, overly rigid information structures can hinder the collection of information while overly
ambiguous structures can draw information into a virtual black hole.

Recently, a new instrument has appeared in the knowledge manager’s toolbox—folksonomy. Folksonomies represent a nearly diametrically opposite approach to traditional information organization. For example, while ontologies rely on knowledge management experts to develop specific functional definitions, folksonomies impose no preconceived definitions and allow group consensus to reinforce appropriate classifications that emerge organically.

Hierarchies attempt to organize information and give context to data through a branching structure while folksonomies allow for a multiplicity of contexts. Rather than working from the top down to build a structure and then insert data, folksonomies start at the data level and allow communities of knowledge for consumers to apply their own organization, in the form of tags and metadata, to whatever information they see as valuable. “Folksonomies have allowed communities of users to build structure on top of content using tags as annotations” (Dubinko, Kumar, Magnani, Novak, Raghavan, & Tomkins, 2006). As metadata grow, the context of the information also broadens and thickens. This process leads to folksonomies adapting to their communities’ needs and offers a flexible strategy for maintaining dynamic information resources: “[S]ystems employing free-form tagging that are encouraging users to organize information in their own ways are supremely responsive to user needs and vocabularies, and involve the users of information actively in the organizational system” (Mathes, 2004).

One of the most frequently observed challenges of the use of folksonomies in knowledge management is the lack of control and structure around the use of tags (Peters, 2006): “A folksonomy represents simultaneously some of the best and worst in the organization of information. Its uncontrolled nature is fundamentally chaotic, suffers from problems of imprecision and ambiguity…” (Mathes, 2004). However, it remains to be determined whether, in scaling down a folksonomy from its traditional global scope (involving millions of contributors) to a size more commensurate with corporate-sized KM solutions, there is enough metadata to make a folksonomy effective.

Large-scale social bookmarking sites (such as Del.icio.us or Reddit.com) have been among the earliest adopters of using folksonomies to organize information. These sites have been drawing a great deal of attention from the web-surfing public and some now claim to have client bases in the millions (Del.icio.us, n.d.). There is a plethora of anecdotal evidence to suggest that these sites are both appreciated and effective tools for storing, finding, and sharing Internet-based resources—a form of social knowledge management. Much of the success of social bookmarking is attributed to its loosely structured approach to organizing data and the ease with which consumers can learn and integrate a folksonomical strategy. If this knowledge management strategy could be adapted to a smaller scale, one that more closely matches user needs in a business environment, the benefits to the knowledge management industry would be significant. As Mathes (2004) notes, “Folksonomies leverage the expertise of knowledge consumers to create an opportunity that has the lowest investment in time effort and cognitive cost to collect the information.”

The purpose of this chapter is to examine the critical success factors in the development of a small-scale folksonomy and outline the key features needed to build a folksonomy that can be scaled to a size that might be more readily operationalized in organizations.

BACKGROUND

Folksonomies can be described as ad hoc information management systems that acquire their entire structure through the context created by the descriptive contributions from an interested
Critical Success Factors in the Development of Folksonomy-Based Knowledge Management Tools

Figure 1. Entity relationships of a folksonomy

![Diagram showing the relationship between actor, tag, and object in a folksonomy]

and involved community (Brooks & Montanez, 2006). Within a specific folksonomy, there is a requirement that three entities exist: an object being classified; at least one descriptive single word tag to categorize the object; and an actor to interact with the object. For example, Jordan (the actor) may tag a photograph of a new puppy (the object) as “photo, puppy, Kujo, cute” (four distinct tags). While this example only manages data from a single knowledge consumer, theoretically other actors, such as Jordan’s family, could make contributions. Regardless of how information is added to the folksonomy, it is important to recognize that a tag reflects the unique perspective of the contributor. Although two actors may use the same tag on differing objects, there may be disagreement between the actors as to the specific meaning of that tag. The argument therefore can be made that tag selection represents how knowledge consumers think about data rather than how the data owners/managers perceive and classify it (Fichter, 2006). For this reason, folksonomies need to associate an object and its tags with the contributor to maintain a full definition of the object. Figure 1 illustrates this relationship.

Although a knowledge management solution is meant to facilitate the sharing of information, the motivation for individuals to contribute to such a system is somewhat more selfish (Fay, 2007). By allowing members to have their own personal lists of contributions, the system grows around the needs of the knowledge consumer. In this way, it is the uncoordinated and selfish interests of the community’s individual members that create the folksonomy (Cattuto, Loreto, & Pietronero, 2006). Each contributor adds links to a bookmarking site and tags it in a way that will facilitate future use. For an individual, the focus of activity is on personal needs and not the needs of the community. The community benefits from the activity by aggregating the selfish actions of all the participants: “Folksonomy is the result of personal free tagging of information and objects (anything with a URL) for one’s own retrieval. The tagging is done in a social environment (usually shared and open to others). Folksonomy is created from the act of tagging by the person consuming the information” (Vander Wal, 2007).

Critical Success Factors in Building a Folksonomy-Based Knowledge Management Tool

The popular trade media have already pointed at a number of social networking sites and enumerated several common characteristics for a folksonomy-based knowledge management tool. Importing data, popular links, recent additions, friends, ratings, privacy, tag suggestions, RSS, widgets, browser add-ons, among others, were identified by Read Write Web (Alex, 2006). The editors of the “Lifehacker Blog” also emphasized the importance of ranking, RSS, and various forms of collaboration as common features (Lifehacker, 2007). Consultants Common, a social bookmarking comparison tool, indicated several key factors that included: RSS, tagging, descriptions, bookmarklets, tag browsing, sharing, importing, exporting, integration, and page caching (Webb, Kanter, White, Simon, & Loving, 2005). Therefore, the following critical success factors can be identified as being relevant to small-scale folksonomies:
Critical Success Factors in the Development of Folksonomy-Based Knowledge Management Tools

- Personalized reference lists
- Community lists
- Tag clouds
- Metadata
- Tag recommendations
- A searchable database
- Cross-linked tags, and
- Cross-linked users.

If an organizationally appropriate KM system were to be built based on the experiences of large-scale social bookmarking sites, the evidence cited suggests that these eight features need to be included in the design.

Operationalizing the Critical Success Factors

For folksonomies to be adopted by knowledge workers in substitution for a traditional and hierarchically structured KM system, there needs to be an exploration of some of the limitations or boundaries that constrain the folksonomical approach. Adopters must determine whether a folksonomy would work with a small number of contributors and how its performance would be measured. Three lines of inquiry can be identified that need to be explored to evaluate a folksonomy’s effectiveness as an organizational knowledge management tool:

- Identification of empirical measures of user interaction
- Identification of key features that would entice participation in the use of such a folksonomy tool, and
- Development of a user satisfaction instrument to validate findings based on the first two challenges.

In effect, adopters must ask what a folksonomy might look like when implemented, what would people do with the tool, and how well does it satisfy their needs? “Performance measurement,” a measure of user interaction, is a crucial measurement in knowledge management; in this regard, log files tracking user behaviour play a key role. Zipf’s principle of least effort developed in “Human Behaviour and The Principle of Least Effort” (1949), also known as Zipf’s Law, has been shown to be a signature behaviour for self-organizing human activity (Cattuto, Loreto, & Pietronero, 2006). Zipf argued that people are predisposed to seek solutions that require the least effort; in essence, people create tools to solve problems more easily. The principle of least effort states that an information-seeking actor (person) will tend to use the most convenient search method available, generally opting for using that search method’s least exacting mode. In other words, if an online search tool is most readily available, as opposed to a printed resource, such as a textbook, the person will tend to search online. However, if the textbook is more useful (or the computer is further away and not turned on, for example), the person will use the textbook. Information-seeking behaviour stops as soon as minimally acceptable results are found. Zipf’s principle holds true regardless of the user’s proficiency as a searcher or their level of subject expertise; it takes into account the user’s previous information seeking experience. The user will use the tools that are most familiar and easy to use that finds results. The principle of least effort is known as a deterministic description of human behaviour. The principle applies to any information seeking activity. Zipf’s argument is further supported by the notion of the selfish user (Fay, 2007).

Additionally, one can use Zipf’s Law to understand how the principle of least effort is supported by the use of tags and tagging behaviour. Zipf’s Law was developed using empirical data and statistical procedures; it refers to the fact that many types of data can be approximated with a Zipfian distribution, one of a family of related discrete power law probability distributions. Zipf’s Law maintains that given a body of language occurrences, such as tags, the frequency of any word
Critical Success Factors in the Development of Folksonomy-Based Knowledge Management Tools

in a body of text is inversely proportional to that word’s rank in a frequency table. Thus the most frequent word will occur approximately twice as often as the second most frequent word, which occurs twice as often as the fourth most frequent word, and so on.

Since large-scale social bookmarking sites can be viewed as following Zipf’s principle of least effort, it can be argued that if a small-scale folksonomy also follows the principle of least effort, the information held within the folksonomy must have similar value to the community. When folksonomies are deconstructed and tested, they are shown to follow the principle of least effort (Zipf, 1949). While there are a number of distributions associated with tagging and folksonomies (Peters, 2007), Zipf’s work clearly draws a connection between a power distribution and the knowledge consumers desire to act selfishly and reduce the effort that they are required to invest (Zipf, 1949). Butler (2004) succinctly sums up the connection: “I think the lack of hierarchy, synonym control and semantic precision are precisely why it works. Free typing loose associations is just a lot easier than making a decision about the degree of match to a pre-defined category (especially hierarchical ones).”

Within a folksonomy, there are certain actions that can always represent the transfer of information. These value-indicating actions are behaviours identified in the system logs where participants are either accessing a resource listed in the database or contributing to the database in some way: adding a link to a resource; adding a new resource; or editing the tags and metadata associated with an existing object. Therefore, the first line of inquiry - the identification of empirical measures of user interaction – can be addressed through examining log files for the types of value-indicating actions.

When displaying a collection of tags, social bookmarking sites often use minimal formatting; this gives a tag collection a chaotic appearance. To help identify which tags are most common, individual tags are often emphasized using a combination of font sizes and colors. This gives the collection of information a cloud-like appearance and has lead to the adoption of the term tag cloud. Tag clouds appear to serve two purposes: (1) they display all the tags associated with an object, thus allowing the knowledge consumer an opportunity to holistically evaluate the community’s perception of an object; and, (2) they allow for a sense of key themes related to a given object. Figure 2 shows an example of a tag cloud with varied font sizes and clickable links.

Folksonomies derive their structure from the reinforcement of tags through repeated use. If a tag referring to an object is used more frequently than another, it can be argued that that tag is perceived as more important to the description of the object by the users of the folksonomy. To create reinforcement behaviour, social bookmarking sites often display some or all of the tags that have already been used to describe an object. This reinforcement behaviour, combined with the development of a tag cloud, satisfies several of the critical success factors listed earlier.

Although folksonomies are popularly associated with the use of tags, social bookmarking
Critical Success Factors in the Development of Folksonomy-Based Knowledge Management Tools

sites frequently included additional metadata in the form of a comments field. This additional field is also tied to the search engine feature built into many social bookmarking sites. Social bookmarking is seen as community activity. Not only can tags be used to describe objects they can also be used to cross-link other objects, tags and users into common themes. This cross-linking ultimately exposes knowledge consumers to broader definitions for tags and could possibly impact how individuals select tags.

The final key to the measurement puzzle is related to user satisfaction and whether a tool, such as a folksonomy, actually improves participation in knowledge management projects. A simple survey tool can be used to explore such questions as:

- How satisfied were you using a tag cloud?
- How satisfied were you using a search tool based on metadata?
- How satisfied were you using your own tag lists to retrieve material?

Further, questions exploring whether tags conveyed important information, the degree of control a person felt over the data, and how effective the knowledge consumer felt the tool to be when coupled with the secondary data collected from log files will lead to a clear understanding of how well a folksonomy can be adapted to corporate or small-scale knowledge management.

Testing the Critical Success Factors

An online experiment was conducted at Lakehead University, where a Learning Object Repository was constructed based on the success factors identified above. The repository incorporated all eight critical success factors in its design: personalized reference lists, community or global lists, tag clouds, metadata, tag recommendations, a searchable database, cross linked tags, and cross linked users. Post-secondary instructors in Management Information Systems were invited to contribute to the repository and use a folksonomy to organize their materials. Forty-eight teaching professionals and other interested parties were tracked as they used the system. In two months, the participants were surveyed on their impressions of the Learning Object Repository.

To construct an effective small-scale folksonomy, a three-pronged approach was developed. The first step was to scrutinize the public social bookmarking environment and, as described, eight critical success factors were identified as crucial to developing a working folksonomy based tool. Next, metrics for evaluating and assessing user engagement were developed using information from log files. Although many important characteristics of performance can be derived from log files, they are not the sole source of information regarding user satisfaction. As discussed above, only direct questioning of a system user can address satisfaction issues.

All web hits entering and exiting the site were tracked to identify value-indicating actions. At the end of the trial period, the repository had been visited a total of 467 times. The repository grew to 37 objects with 172 distinct tags being used 299 times. During the experiment, the website was visited an average of 8.19 times every day and 490 value-indicating activities occurred in 203 of the visits to the repository. Of the 490 value-indicating activities recorded, 86% of those activities involved following links to resources within the Learning Object Repository. When tags within the system were sorted by popularity and their frequency plotted against their rank, the results fully accord with the predictions consistent with Zipf’s Law.

The survey distributed to all the participants at the end of the research period had a 29% response rate. The results indicated that participants considered the community tag cloud useful, scoring it 7 out of 10. The survey also found that satisfaction with an individual’s own tag cloud was slightly more conservative, scoring 5.69 out of 10. The
study determined that the folksonomy site had been an effective tool for managing information within small groups and that the critical success factors for a folksonomy were as described.

**FUTURE TRENDS**

Future applications of folksonomy-based approaches to knowledge management will require designers and managers to see past the current drawbacks that arise from the use of an uncontrolled vocabulary in the tagging activities. Indeed, it is this non-predefined use of language that presents a number of opportunities and unique benefits.

Using folksonomies to sort and organize images, for example, appears to be a good match to the capabilities of this KM tool. As others have noted (e.g., Jörgensen, 1999; Gordon, 2001), the classification of images using standard KM hierarchies is both challenging and time consuming and may not lead to a usable data set. On the other hand, browsing as a method of navigating an image collection is a fitting solution and within the capabilities of a folksonomy. Tagging images with user-defined tag sets encourages browsing and may add an element of serendipity to the results, which increases the value of the search results.

A major benefit of the folksonomical approach is that it begins with a blank slate on which the structure of a content space can be allowed to develop through use until patterns emerge (Haverty, 2002). Giving the end user control over the organization of content can also allow entirely new content domains to develop.

It becomes possible, then, to create specific tags to allow them to share new content forms that either have not been conceptualized at the time a hierarchical KM tool was designed or that are extremely difficult to classify under a hierarchical structure. For example, participants in Flickr.com have developed tags that permit them to share images of their computer desktops or instant messaging status as a sort of personal history or discovered narrative. In this way, categories become things, and the classification becomes a shared space for communities of users to explore and develop (Lakoff, 1987). As Herwig notes: “One of the outcomes of the research into folksonomies is FolkRank, a search algorithm that exploits the structure of folksonomies; the name reveals that it was inspired by PageRank, but as the graph of folksonomy structures does not correspond to the web graph, some adaptations had to be made” (2008).

Other trends derived from this approach include: detecting trends (in tag sets); detecting spam in tag sets (a major challenge) to improve the results; and developing a logsonomy (i.e., the structure of search engine query log files). Dawson (2007) proposes five key frames for leveraging knowledge in organizations that are emerging as the successors to knowledge management and that executives find relevant, compelling, and actionable: social networks, collaboration, relevance, workflow, and knowledge-based relationships.

**CONCLUSION**

Based on the results of the above study, it can be concluded that the critical success factors identified above are in fact effective design elements in a folksonomy-based knowledge management tool. The critical success factors are: personalized reference lists, community lists, tag clouds, metadata, tag recommendations, a searchable database, cross linked tags and cross linked users. Best design practices were identified from a selection of popular social networking sites and criteria for measuring the effectiveness of a folksonomy and user satisfaction were established.

One of the major benefits of the folksonomical approach to knowledge management, as has been noted in this chapter, is the fact that it is developed and maintained by the users of that body of knowledge. In much the same way that wiki tech-
nology improves the availability of knowledge, folksonomies make the storage and dissemination of knowledge much more accessible to the user community itself. Although a folksonomy-based KM system is less expensive, more flexible (scalable), and user-friendlier, a socially created folksonomy demonstrates strength and value. Communities of users create storehouses of shared meanings and understanding when they contribute to the tag cloud of any given folksonomy. As this storehouse of knowledge evolves over time, it has the dual outcome of creating a more viable knowledge management tool while strengthening the bonds of the user community, which facilitates the longevity of the tool.

REFERENCES


**KEY TERMS AND DEFINITIONS**

**Folksonomy:** Folksonomies can be described as an ad hoc information management structure that acquires its entire structure through the context created by descriptive contributions from an interested community. A folksonomy is a user-created taxonomy used to categorize information. Folksonomies rely on single word tags assigned by users to create context and establish the value of information.

**Hierarchies:** A series of ordered groupings within a system, moving from broad categories to narrowly defined.

**Knowledge Management:** Knowledge management is the process by which information is collected and disseminated for the purpose of
improving the effectiveness of a decision making process.

**Metadata:** Metadata are structured data that describe the characteristics of a particular resource.

**Ontology:** The goal of an ontology is to ensure that all users of a given set of data have the same understanding of the words and concepts within that dataset. In knowledge management, ontologies focus on creating dictionaries of commonly defined terms.

**Semantic Web:** This term refers to the development of a highly structured XML-based WWW alternative that is more conducive to automated analysis.

**Social Networks:** The term refers to individuals and their interconnections with others, which could include business relationships, kinship or trust.