Benefiting from the hive: Folksonomy in technical libraries

Nader Naghshineh
Information Studies Lab, University of Tehran
NNaghsh@ut.ac.ir

Mahshid Eltemasi
Library and Information Science Dept., University of Tehran
elt.mah@gmail.com

Abstracts:
Internet librarianship, Library 2.0 and Webinomics are but few examples of the terms that have come to affect our Net-Centric existence. Cooperative organization of web-based resources or Folksonomy is yet another facet of this existence. While per se it deals with electronic based or Net-based materials, it is not totally in keeping with the fact that it does not cover the physical libraries. The present paper presents the experience gained in using Folksonomy and establishing a Library 2.0 system in a technical University Library. Using a variation of Referral Technical Research Service for a limited period, a 30 day snapshot of query language employed by the patrons was developed. A team poured over the data to analyze both the queries and the results in order to discern patterns. The pattern(s) were used to optimize tagging and services. The patrons’ intrinsic tagging behavior was incorporated into a collaborative tagging routine. Using a cognitive construct to better understand tagging relegation by the patrons, this tagging was extended to tangible library information assets. In essence, two levels of access was incorporated, one level being static and based on library practice, and the other being a level dynamically shifting based on the customer profile and cognitive style.

Preface
Every Library, great and small, whether in major urban centers or in the back mountains of Zagros, use a form of classification for bringing about order to their collections. The running joke here in Iran is that some Librarians take this to an extreme of even classifying their own personal life. However, the common classification schemes are still based on and aimed at printed sources. There is a degree of physicality inherent in this statement. Every classification aims at tagging a printed, tangible source. But the internet resources or what is often termed as Soft-Sources do not often lend themselves easily to such classification. Yet we still maintain that the primary mission of classification and cataloging is to deliver the appropriate information item, on time to the appropriate client.

We are also faced with the reality that most of our client are predisposed to satisfy their information queries through internet, databases and electronic resources. In fact, there is an ISL test whereby the LIS graduate students are required to mine for information without the benefit of their laptops and UT Digital library. Known as the E-deprivation, the test is part of students’ training for cognitive flexibility.

There is also a new possibility accorded to libraries via social softwares. It is now possible to turn Libraries into collaborative ventures, through implementation of various community-building tools. The stakeholders would redefine the library holdings and importance within their own frames of reference.

Folksonomy (also known as social tagging, social indexing, social classification, and collaborative tagging) is the practice and method of collaboratively generating and managing tags to annotate and categorize content. Metadata is generated not only by experts but also by creators and consumers of the content. Usually, freely chosen keywords are used instead of a controlled vocabulary (folksonomy,2008).

What follows is a brief description of an attempt made by a group of librarians in gaining practical experience with folksonomies within a library environment.

Project Description
The Islamic Azad University has the widest geographical and ethnic distribution among the Iranian Universities. It enrolls close to 1.3 million undergraduate and graduate students spread over 30 states. While the quality of service is not homogenous, the libraries located near major metropolitan areas, especially in Metro Tehran display adequate service levels.
One of these libraries, located 80 miles Southwest Tehran, has established an information support unit where a number of librarians mediate information search and delivery for both students and the faculty. The search is carried out over internet resources and the databases subscribed by the Islamic Azad University. In compliance with the information counselor protocols (Naghshineh 2003), one counselor is assigned to service several fields using a dedicated station. Every station stores and keeps records of all the phrases and keywords searched for a 30 days period. The departments served were Industrial design, Architecture, Electronic engineering, Computer Science and Mechanical Engineering.

Following the 30 days period, all searches with the resources retrieved were analyzed and data-mined. Visualization techniques were also employed to assist pattern discovery. In effect the users would have tagged their required information upon retrieval, thus contributing to a useful database of tagging commonality. In reality, the keywords suggested by both students and the faculty were free formed and were stored without any vocabulary control. The essence of Folksonomy is the combined wisdom of web users in collection, organization, management and dissemination of information. However, as part of knowledge management strategy the combination of these two different tagging practice would point out explicit islands in the tacit seas (Hicks, Dattero and Galup 2007).

In effect the research aimed at finding a low tech means of establishing a bridge between two systems that provided a means or method for making automatic prediction about the interest of a user. This is often done in marketing and in certain sub-domains of psychology by gathering information about the likes or dislikes, or taste, from other users. This mode of automatic prediction is also a form of filtering known as collaborative filtering (Goldberg, et al. 1992). The basic notion of such approach is that those who concurred in the past are likely to agree again in the future. The recommender system we often come to see in sites such as Amazon and some digital libraries are in effect one such collaborative filtering (CF) system.

Interestingly enough it seems also that the information is valued as long as it has been already filtered (or in internet parlance, ranked and rated) by other people (Origgi 2008). One aspect of such ranking system when viewed in the light of Folksonomic analysis could be the mapping of the cognitive styles employed in such rating system.

**Project Findings**

Following a 30 days period, a snapshot of the keywords and phrases was made. The following results were obtained:

1. There were a large number of common or redundant keywords. When the frequency of a word is high, there is a high degree of certainty that it would be accepted as a main tag
2. Different people offered different versions of a particular word. This is especially important when you deal with Persian Script whereby you can combine or separate the suffixes. This lead to serious thinking about the kind of recommender system envisioned for the library.
3. Homonyms and synonyms provided in farsi as well as pidgin argot posed the biggest challenge
4. Lack of any subject hierarchy. It seemed that the users considered all subjects to share the same value in ranking and rated them as such.
5. With respect to information resources such as images and drawings, instead of articulating search terms, users often express their emotions. There seemed to be no consensus on categorization of such emotives..

In order to surmount the issues confronting the project, it was decided that instead of creating a alphabetical list of tags presented by the users and then linking the relevant resources to this list, the information counselors make a limited intervention. In what that, for the lack of any better terms, could be described as a controlled folksonomy, the following actions were taken:

- A consensus was reached to adopt one term from among the similar tags offered. The information Counselors made effort to have more diversified information resources available on the subject.
- While acknowledging language variances only one form was used in the tag (applicable only to Persian Language)
• Homonyms were identified
• Subject Specialists were recruited to overcome the issues of semantic as well as misspelling mistakes.
• A shifting thesaurus function for the tags was envisioned.
• A/V sources were given a folksonomic tags

The library was thus able to develop an internal, web-based thesaurus or tag book. Users accessing the library webpage would encounter a page containing a list of tags that they had determined. This tag list has been reviewed by experts. The tag list is supported by the sources retrieved from databases, the electronic resources and even print resources held by the library. Thus when a user clicks on one of the entries on the list, another page would open that contains a summary of the database articles with pertinent links, bibliographic information on the library books along with retrieval and ranking codes. The list would also indicate the availability of non-print materials.

References


