# Controlled Vocabulary and Folksonomies

Louise Spiteri

School of Information Management

DALHOUSIE UNIVERSITY
*Inspiring Minds*

# What are folksonomies?

- Folksonomies (known also as "social classifications") are user created metadata.

- They are grassroots community classification of digital assets.

- The term "folksonomy" was created by Thomas Vander Val and represents a merging of the terms "folk" and "taxonomy."

- One form of explicit user created metadata was popularized in the late 1990s with link-focused websites called weblogs.

# Where are folksonomies found?

- Folksonomies are found in social bookmarks managers such as Del.icio.us (http://del.icio.us/) and Furl (http://www.furl.net/), which allow users to:
  - Add bookmarks of sites they like to their personal collections of links
  - Organize and categorize these sites by adding their own terms, or tags
  - Share this collection with other people with the same interests.

- The tags are used to collocate bookmarks: (a) within a user's collection; and (b) across the entire system, e.g., the page http://del.icio.us/tag/blogging will show all bookmarks that are tagged with "blogging" by any user.

# Inter-term relationships

- There are no clearly defined relations between and among the terms in the vocabulary, unlike formal taxonomies and classification schemes, where there are multiple kinds of explicit relationships (e.g., broader, narrower, and related terms) between and among terms.

- Folksonomies are simply the set of terms that a group of users tagged content with; they are not a predetermined set of classification terms or labels.

# Popular folksonomy sites

- [Del.icio.us](http://del.icio.us) (http://del.icio.us)
- [Flickr](http://www.flickr.com) (http://www.flickr.com)
- [Frassle](http://www.frassle.org) (http://www.frassle.org)
- [Furl](http://www.furl.net) (http://www.furl.net)
- [Simpy](http://www.simpy.com) (http://www.simpy.com)
- [Spurl](http://www.spurl.com) (http://www.spurl.com)
- [Technorati](http://www.technorati.com) (http://www.technorati.com)

# How folksonomies work

- Registration is free. Little personal information is required; normally just a login name and password.

- Once registered, you add a bookmarklet to your browser. When you find a web page you'd like to add to your list, you add it to your manager site. You then assign keywords to describe the content of the site.

- If your page has been bookmarked by other people, you will be shown the most popular tags assigned to it; you can assign your own tags, or simply click on the popular tags to have them assigned automatically.
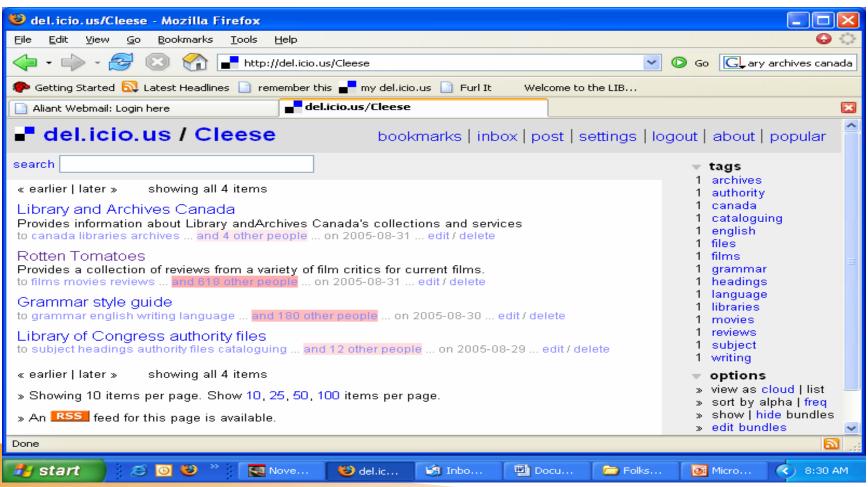
# Bookmarklets on browser toolbar

# Adding tags to bookmarks

# Examples of my bookmarks

# Other people who share one of my bookmarks

# The popularity of folksonomies

- The growing popularity of folksonomies can be attributed to two principal factors:
  - An increasing need to exert control over the mass of digital information that we accumulate on a daily basis.
  - A desire to "democratize" the way in which digital information is described and organized by using categories and terminology that reflect the views and needs of the actual end-users, rather than those of an external organization or body.

# Bookmarking

- One of the most daunting challenges of information management in the digital world is the ability to keep, or re-find relevant information.

- Bookmarking is one of the most popular methods for storing relevant Web information for re-access and re-use.

- Many of us have folders within folders within folders. We find ourselves bookmarking the same site a dozen times because we can't remember where we filed it.

- Bookmarks are often stored in the browsers of the different computers that we use.

# Managing bookmarks

- Social bookmark managers such as Del.icio.us allow people to store their bookmarks centrally on the bookmark managers.

- Because the bookmarks are on the Web rather than within a single instance of a browser, they can be accessed from any computer with an Internet connection.

- The bookmark managers assume that once users find a good site, they not only want to share it with others, but want to discover other related sites and people who are interested in the same topic.

# Folksonomies and user vocabulary

- In information retrieval systems (IRS), the vocabulary used to organize content may be based upon the choices of the authors of the materials, the designer of the IRS, or the designer of the controlled vocabulary in place.

- Folksonomies reflect users' choices in diction, terminology, and precision.

- Folksonomies can adapt very quickly to changes in user needs and vocabulary, and adding new terms to a folksonomy incurs virtually no cost for either the user or the system.

# Folksonomies and online communities

- Folksonomies create a sense of community amongst their users. Most social bookmark managers will recommend new links and other members' folders or sites that are strongly related to an individual member by analyzing his or her linking pattern.

- As soon as users assign a tag to an item, they can see the cluster of items carrying the same tag. This feedback loop leads to a form of asymmetrical communication between users through metadata.

- The users of a system negotiate the meaning of the terms in the folksonomy.

# Ambiguity

- The terms in a folksonomy may have inherent ambiguity as different users apply terms to documents in different ways.
  - E.g., the tag "ANT" has been used to refer to "Actor Network Theory", a sociological term, as well as Apache Ant, a Java programming tool

DALHOUSIE
UNIVERSITY
*Inspiring Minds*

# Polysemy

- The polysemous tag "port" could refer to a sweet fortified wine, a porthole, a place for loading and unloading ships, the left-hand side of a ship or aircraft, or a channel endpoint in a communications system.

# Synonyms

- Folksonomies provide for no synonym control; the terms "mac", "macintosh", and "apple", for example, are used to describe Apple Macintosh computers.

- Both singular and plural forms of terms appear (e.g., flower and flowers), thus creating a number of redundant headings.

# Specificity

- Related terms that describe an item vary along a continuum of specificity ranging from very general to very specific; so, for example, documents tagged "perl" and "javascript" may be too specific for some users, while a document tagged "programming" may be too general for others.

# Syntax

- Folksonomies provide no guidelines for the use of compound headings, punctuation, word order, and so forth; for example, should one use the tag "vegan cooking" or "cooking, vegan"?

DALHOUSIE
UNIVERSITY
*Inspiring Minds*

# Incorrect usage

- Tags could be applied incorrectly; the term "archeology", for example, is used to tag items pertaining to both dinosaurs and primitive microbes

# Consensus

- Users strive to achieve a degree of consensus over the general meaning of tags.

- As a URL receives more and more bookmarks, the set of tags used in those bookmarks becomes stable across different users.

- This stabilization is facilitated through imitation and shared knowledge. Del.icio.us shows users the tags most commonly used by others who bookmarked the same URL already; users can easily select those tags for use in their own bookmarks, thus imitating the choices of previous users.

# Folksonomies and controlled vocabularies

- Folksonomies are not necessarily antithetical to controlled vocabularies.

- Once you have a preliminary system in place, you can use the most common tags to develop a controlled vocabulary that truly speaks the users' language
    - E.g., you can link related tags such as "nyc," "newyork," and "newyorkcity"; it may be possible to align these terms with established controlled vocabularies, such as the Getty Thesaurus of Geographic Names, in order to provide a greater range of related terms.

# Other uses for folksonomies

- Could be used to organize resources for an intranet, course collection, etc. For example, I plan on using Del.icio.us in my Indexing course next term. Students will be encouraged to bookmark and organize indexing sites of interest; at the same time, it would be a good opportunity to study tagging behaviour and patterns.

# Other uses for folksonomies

- Could be used to enhance the customizable features of library catalogues. Clients could organize and tag items of interest from the catalogue, as well as external sources (if allowable).
  - Could share these tags and sources with other clients with similar interests. This could lead to a user-directed reader advisory service.

  - Could use folksonomies to supplement existing LCSH vocabulary in the catalogue, e.g., LCSH does not contain terms for the popular film genres "cult", "drama," or "action."

# Conclusions

- Folksonomies are undoubtedly fraught with the problems typical of uncontrolled vocabularies, but their growing popularity suggests that people are interested and motivated in assigning their own metatags to items of interest.

- One cannot help but wonder whether such enthusiasm for metadata would be the same if people were asked to use only prescribed and standardized vocabularies.

# Other areas to explore

- The cognitive and behavioural aspects of folksonomy use:
    - What is the tagging behaviour of people who use folksonomies?
    - Why do people choose the tags they use; what motivates them to modify these tags; how often do they modify them?
    - How are folksonomies used communally?
    - How do folksonomies foster consensus in the use of tags?
    - How does the community affect which tags are used and how?