@toread and Cool: Subjective, Affective and Associative Factors in Tagging

Margaret E. I. Kipp
Palmer School of Library and Information Science
College of Information and Computer Science
Long Island University
margaret.kipp@gmail.com

CAIS/ACSI 2008
Tagging as Collaborative Classification

- Tagging is increasingly examined as a form of collaborative classification
- Multiple studies have examined the consensus shown by frequency graphs
Tagging as User Classification

• Suggested that users of tagging systems are creating a folksonomy or user based taxonomy (Shirky 2005; Hammond et al. 2005)

• Examination of consensus in tagging (Golder and Huberman 2006; Kipp and Campbell ASIST2006, etc.)

• Comparisons of tags and subject headings (Kipp CAIS2006, CAIS2007; Smith SIGCR2007)
Tagging as Personal Information Management

• Frequent comparison (see Shirky 2005, Kipp and Campbell CAIS2006; Kipp IASummit2007, etc.)

• PIM: paper based and electronic studies

• Examine how users organise or arrange their personal/work documents for use

• Also examines differences between current and archival document organisation
Personal Information Management

• Paper based:
  – Malone 1983 - how users organise their desk
  – Kwasnik 1991 - taxonomy of user categories, mainly activities and tasks and projects
    • also see Bewley et al. 1987; Sellen and Harper 2002, Bowker and Star 1999

• Electronic:
  – Jones et al. 2005 - Keeping Found Things Found - KO helps in breaking a topic down into tasks
  – Mas and Hudon CAIS2007 - user classification schemes
Tagging Characteristics

- majority of tags used are subject related or form related -- exactly like traditional classification (also geographic and date/time)
- In all studies, a minority of tags are not subject related:
  - e.g. toread, cool, fun, cais2007
- What is the apparent purpose of these tags?
- Are they only individual or is there any collective useful information here?
Research Questions

• What patterns of user tagging activity emerge on examination of affective or time and task related tags?

• How do users use time and task related tags or affective tags to indicate the value they see in a document?

• What implications do the use of affective or time and task related tags have for the organisation of information?
Methodology

• Data source
  – Del.icio.us, Citeulike (www.citeulike.org), Connotea (www.connotea.org)

• Collection method:
  – Screen scraping (python)

• Data collected
  – all items tagged with a specific list of 83 affective, project and time and task related tags
The Sites

> keep
your favorite websites, music, books, and more
in a place where you can always find them.

> share
your favorites with family, friends, and
colleagues.

> discover
new and interesting things by browsing popular
& related items.

Learn more »

del.icio.us

discuss your favorite websites, music, books, and more in a place where you can always find them. share your favorites with family, friends, and colleagues. discover new and interesting things by browsing popular & related items. learn more »

sign up now

username
password
password again
email

register

A tag is just a word that describes an item saved on del.icio.us. learn more »

the del.icio.us hotlist also see: popular | recent | tags

hot right now

1. Linux.com | CLI Magic: Searching with find save this
first posted by ploka on July 10
linux find tutorial cli search tags

2. Free Science and Video Lectures Online: Even more physics videos and video lectures! save this

3. ThaTypeTrust save this

1 hour ago

1. Strobist: How To: DIY $10 Macro Photo Studio save this
first posted by yeitsinfina on July 10
photography diy lighting tutorial howto tags

2. Run a Secure Web Proxy in 10 Steps - Pure MRZ Community save this

3. Technophilia: Find great podcasts - Lifehacker save this

http://del.icio.us/popular/
Examples of Non Subject Tags

- Affective Tags
  - cool
  - boring
  - exciting
  - important
  - funny
  - strange
  - favorite/favourite

- Time and Task Tags
  - @toread
  - todo, toblog, tovisit
  - @pending
  - todescribe

- Project Tags
  - cais2007, asist2006
  - 519, lis501
  - week7, lecture-7
Data Collected

• all posts tagged with one of 78 specific tags
  – 48 tags were time and task related, 30 were affective tags (project tags were not collected due to the difficulty of locating these)
  – 78 tags were in English
  – 5 tags were in French (lire, alire, @lire, acheter, amusant)
    • non English tags do not yet appear frequently in the popular tag clouds of these sites
  – 203352 posts in total (1831 from CiteULike, 2891 from Connotea and 198630 from Del.icio.us)
General Results

- some time and task or affective tags are very popular
  - cool, fun, funny, toread in main del.icio.us tag cloud
General Results 2

- ToRead and fun are popular tags on all three sites.

- Top 5
  - Citeulike: fun, ToRead, todo, interesting, cool
  - Connotea: fun, ToRead, important, unread, funny
  - Del.icio.us: fun, ToRead, funny, cool, interesting
Time and Task Tags

- many time and task related tags are variations on toread
  - @toread
  - @read
  - readlater
  - unread

- fewer variations for toread are found on citeulike and connotea

- citeulike offers a way to mark interest in reading an article
is the toread tag useful to other users?

- Amazon's recommendation system relies on purchase data and view data as an indicator of interest
- Search engines use hyperlinks as indicators of interest

could a toread tag have a similar function?

could this function like a colleague's e-mail pointing to an article or book?
Affective Tags

- affective tags represent an emotional reaction to an item
  - cool
  - fun
  - strange

- tags such as cool or fun do not appear to add anything to the subject classification of an item

- seem to be poor candidates for search terms for information retrieval... but they indicate interest
Non Subject Tags with Subject Tags

- non subject tags were frequently used with subject related tags
- academic articles on citeulike and connotea were tagged with terms such as fun and cool
- fun was found most often with physics and math, toread with biology
- what do affective tags add to a tag list?
Examples from Citeulike Database Dump

- 320112 - ForensicsAge written in teeth by nuclear tests
  - 'carbon-dating', 'cool', 'forensics'

- 566928 - Action as language in a shared visual space
  - '519', '519-week11', 'action', 'co-location', 'cscw', 'shared-space'

- 983570 - Physics the google way
  - 'education', 'fun', 'google', 'physics', 'search', 'web'
Discussion

- non subject tags are intrinsically time-sensitive
- express response from user not subject of document
- suggest active engagement with the text
- show that user links perceived subject matter to:
  - specific task
  - specific set of interests
  - specific emotional reactions
Discussion 2

- Use of time and task or affective tags shows that tagging expresses a dynamic relationship between users and documents, suggesting possible new ways of modelling information access.

- Research into personal information management systems show users classify by task and project as well as by subject.
Final Thoughts

• What is the effect of personal and subjective terms such as cool, fun and toread in a social bookmarking system?

• What happens when these terms are aggregated?

• Amazon and Google use personal information to generate popularity or relevance indicators, do non subject tags offer any similar advantages?
Thank you!

Questions?