



# Analyzing Folksonomies

*Andreas Hotho*

*Knowledge & Data Engineering Group  
University of Kassel*

# (Research) Areas

---



Data Mining/  
Machine Learning

Information  
Retrieval

Web 2.0

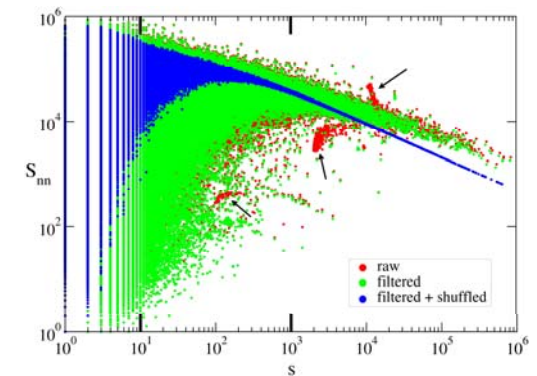
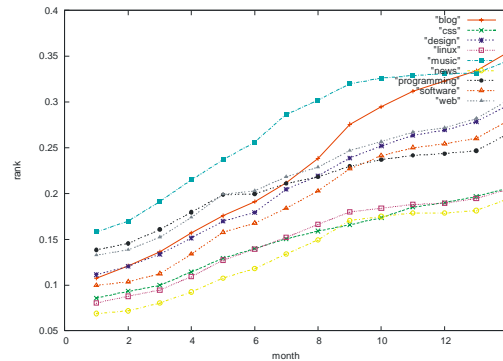
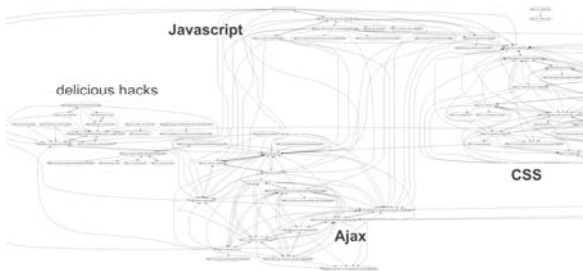
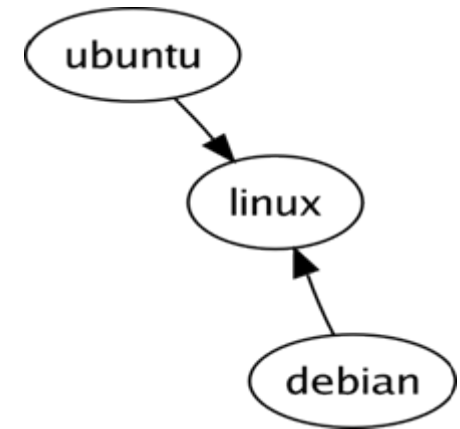
Social network  
analysis

Semantic Web

# Work done in the past

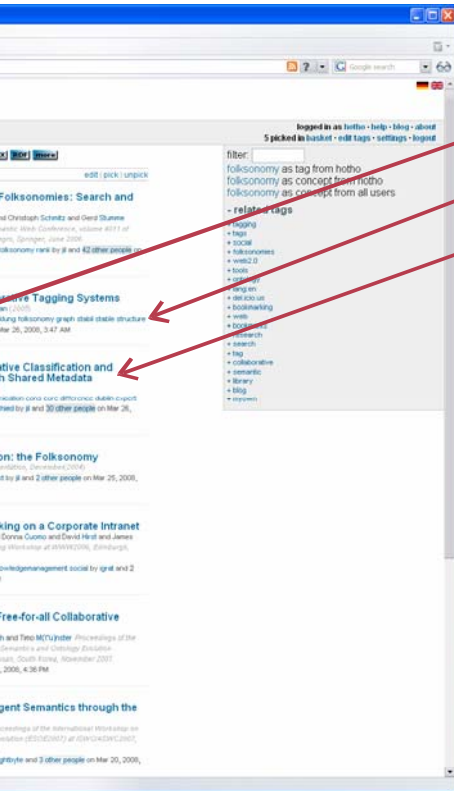


- Ranking in Folksonomies
- Trends in Folksonomies
- Spam
- Tag Recommender
- Learning Ontologies (relatedness measures)
- Community detection
- Analysis of user, tag, resources relationships
- LogSonomy (query log analysis)

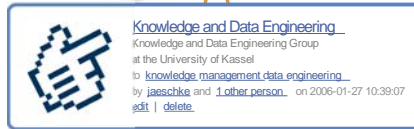




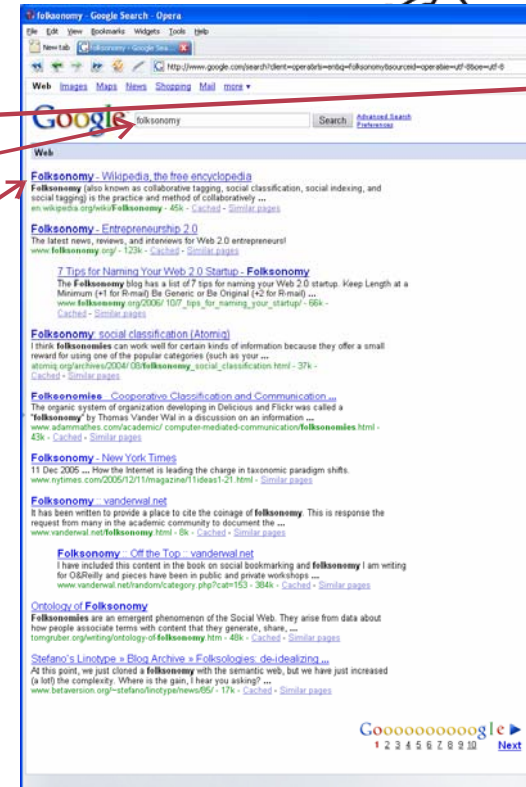
## Folksonomies Logsonomies



Allow **users**  
to assign **tags**  
to **resources**



Allow **users**  
to query **terms**  
and click the **results**



## Social Bookmarking Systems

- Delicious
- Simpy
- BibSonomy

## Query Logs from Search Engines

- Google
- MSN
- AOL



Network measures provide interesting insights into the user behavior of folksonomies and logsonomies

- The degree distribution is not fully powerlaw distributed.
- The logsonomy graph has small world characteristic.
- The cumulative strength distribution is fat tailed.
- The shuffling process destroys semantics in folksonomies as well as in logsonomies (strength/average nearest neighbour strength plot).

Network properties of Folksonomies and Logsonomies (for split queries) are very similar.



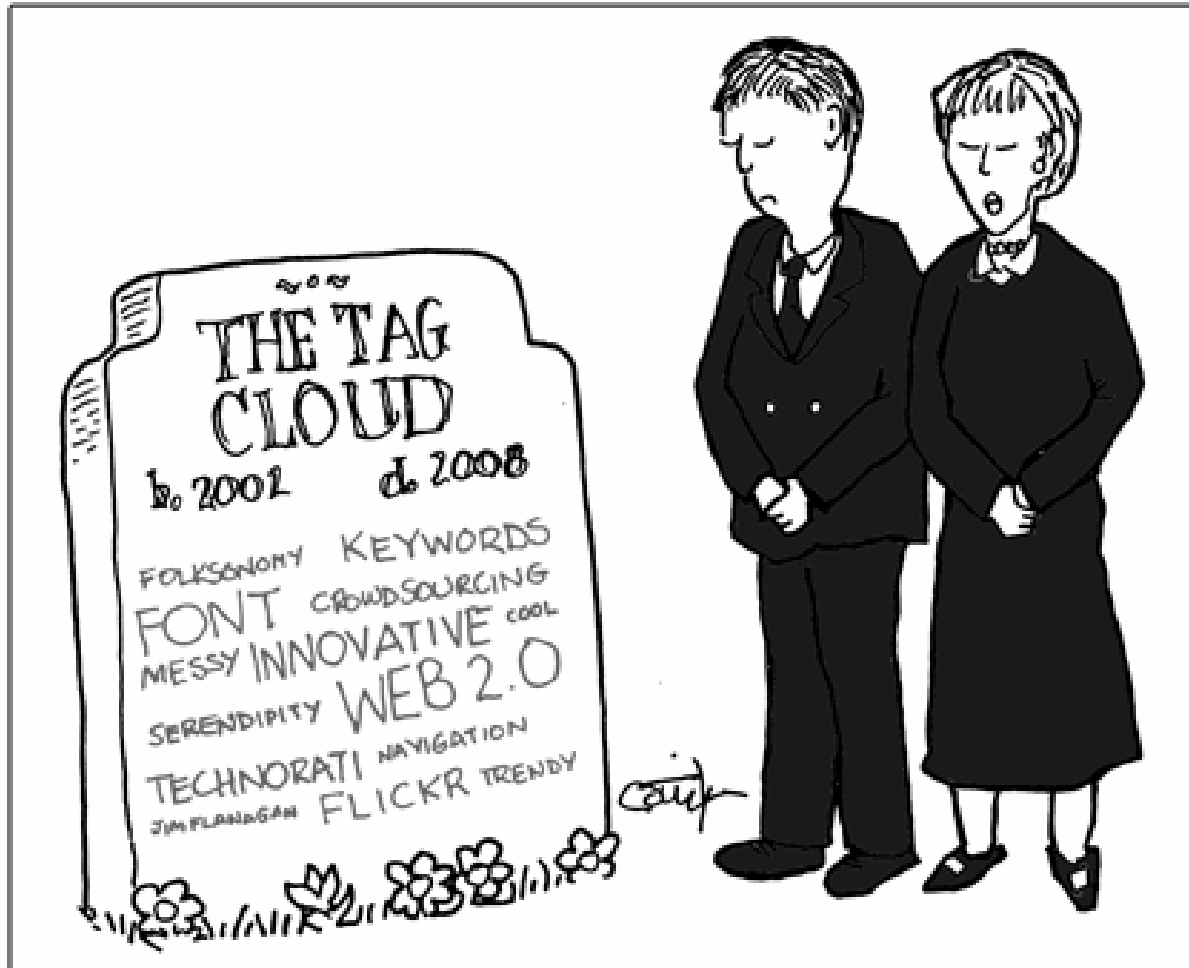
- Merging Logsonomy and Folksonomy for joint ranking.
- Speedup ranking calculation
- Relatedness measures on tags in folksonomies are a good basis to extract semantic relations -> using click data and folksonomy data to extract semantic relationship as well.
- Detecting noise in click data to increase the usefulness of the search engine query log data.
- Detection and Evaluation of Communities in Folksonomies

# The tag cloud is dead long live the tag cloud ;-)



NOISE TO SIGNAL

Rob Cottingham · socialsignal.com/n2s



<http://www.socialsignal.com/image/the-tag-cloud-is-dead-long-live-the-tag-cloud>