



LFE Medieninformatik • Sven Koch

User Interfaces for Machine-Generated Ontologies

**Medieninformatik Oberseminar
Wintersemester 2009/10**





Introduction

What is this work about?

- Tagging Systems (Del.icio.us, Flickr, ...)
- Folksonomies (Folk-Taxonomy)
- Ontologies



Motivation (1/2)

Tagging-systems provide search-interfaces that are sufficient when looking for known, specific items.

Explorative Tasks in Tagging-Systems are currently supported by:

- Search
- Tagclouds



Motivation (2/2)

- Search
 - good as a first step
 - limited if vocabulary is unknown
- Tagcloud
 - easy to understand, but...
 - ... few public terms dominate
 - ... no structure or connections between terms
 - ... few possibilities to support social navigation
- This work
 - Visualize the data of tagging systems without the shortcomings of tagclouds



Contribution & Related Work

- Current work tries to:
 - Provide better navigation in tagging-systems:
 - By using statistical analysis
 - By using external ontological data
 - Derive and maintain ontologies from folksonomies
- This work tries to:
 - Combine statistical analysis and ontological resources to visualize the content of tagging systems



Prior Work

- Clarify technical possibilities / restrictions
 - access to data of tagging systems
 - Flickr (narrow folksonomy)
 - Del.icio.us (broad folksonomy)
 - algorithms
 - distance between resources
 - clustering
 - reduction of high dimensional data
 - access to "ontological" resources
 - OpenCyc ("real" ontology)
 - Google (best coverage)
 - Wikipedia (best trade-off?)



Prior Work

[Prototype Presentation]



Future Work (1/2)

- How to evaluate this work?
 - Automatic evaluation:
 - There is no model solution.
 - There is not a single best solution.
 - User Study:
 - How to objectively measure the results of an explorative task?
 - How to assess users' subjective opinions?

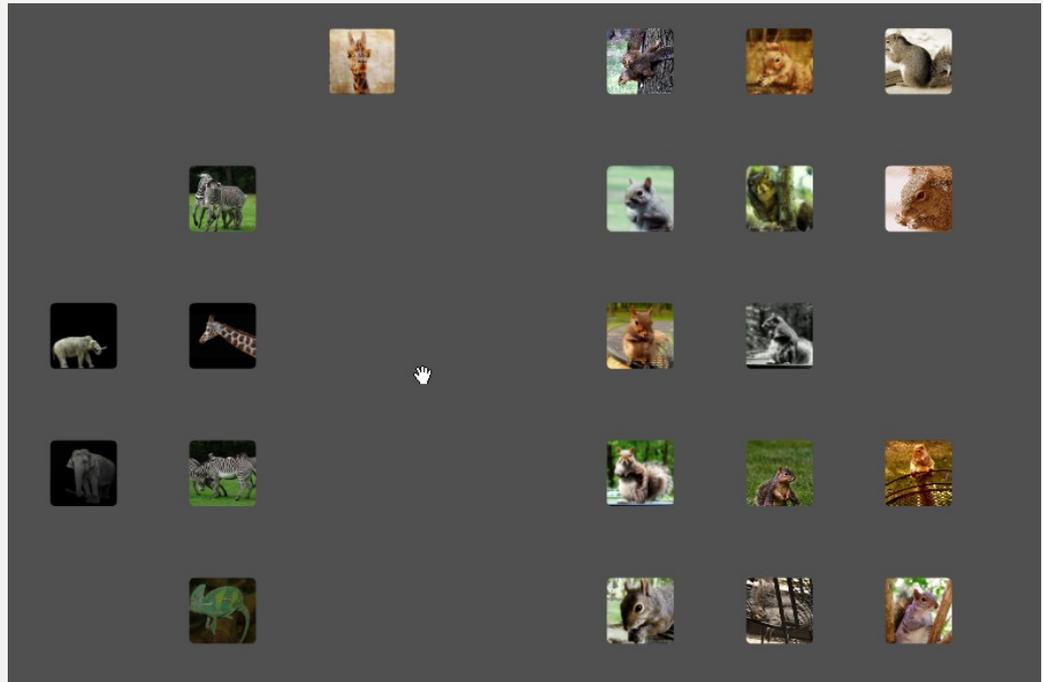
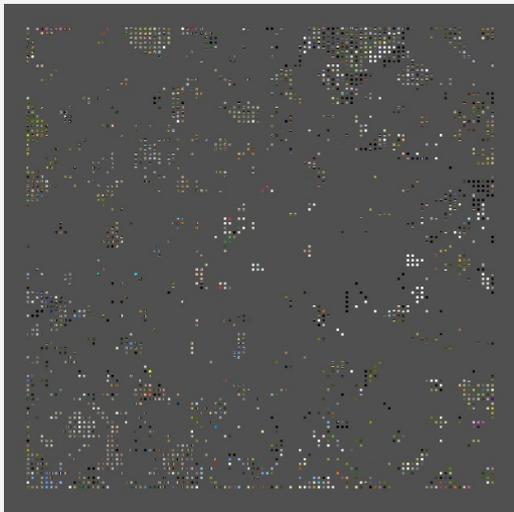


Future Work (2/2)

- Implement functional prototype.
 - Refine current prototype
 - Apply and add techniques to different systems
- Evaluate prototype.
- Iterate (start from 3) as often as possible

Thank you for listening.

Questions?





Appendix – Current Problems (1/2)

- What the user expects: (how can this be known?)

Photo of dog 1
Photo of dog 2
Photo of dog n

Photo of cat 1
Photo of cat 2
Photo of cat n

- What the user sees:

Photo of dog 1
Photo of cat n

Photo of cat 1
Photo of dog 2

- What the system does:

digital color photo of animal taken with canon eos 350d
digital color photo of animal taken with canon eos 350d



Appendix – Current Problems (2/2)

Some numbers of photos in Flickr:

5.961.807 cat

77.994 siamese

53.123 cat siamese (68% of of siamese also tagged cat)
(siamese is probabyl a sub-category of cat)

4.853.659 animal

5.961.807 cat (more cat photos than animal photos)

498.232 animal cat (8% of cat are also tagged animal)

498.232 cat animal (10% of animal are also tagged cat)
(probably no hierarchical connection?)



Appendix – Contribution

