

## Deliver who I mean

Lars Maehlmann

Hamburg, 04.12.07

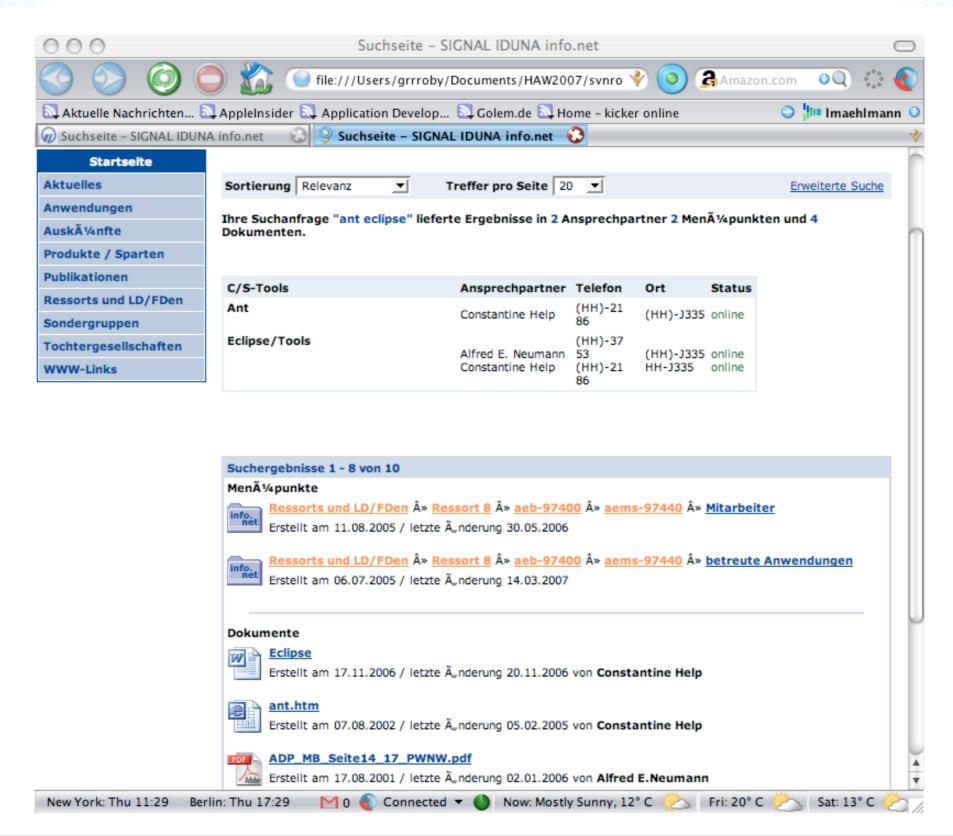


### **A Result**



Hochschule für Angewandte Wissenschaften Hamburg

Hamburg University of Applied Sciences



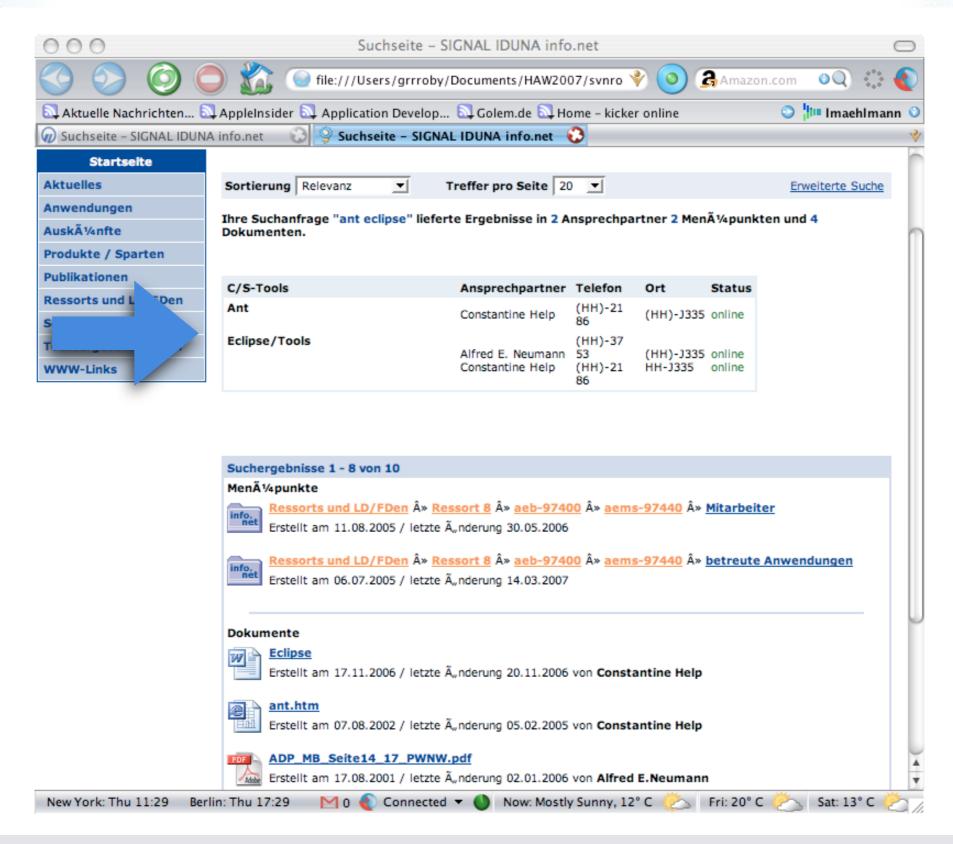


### A Result



Hochschule für Angewandte Wissenschaften Hamburg

Hamburg University of Applied Sciences



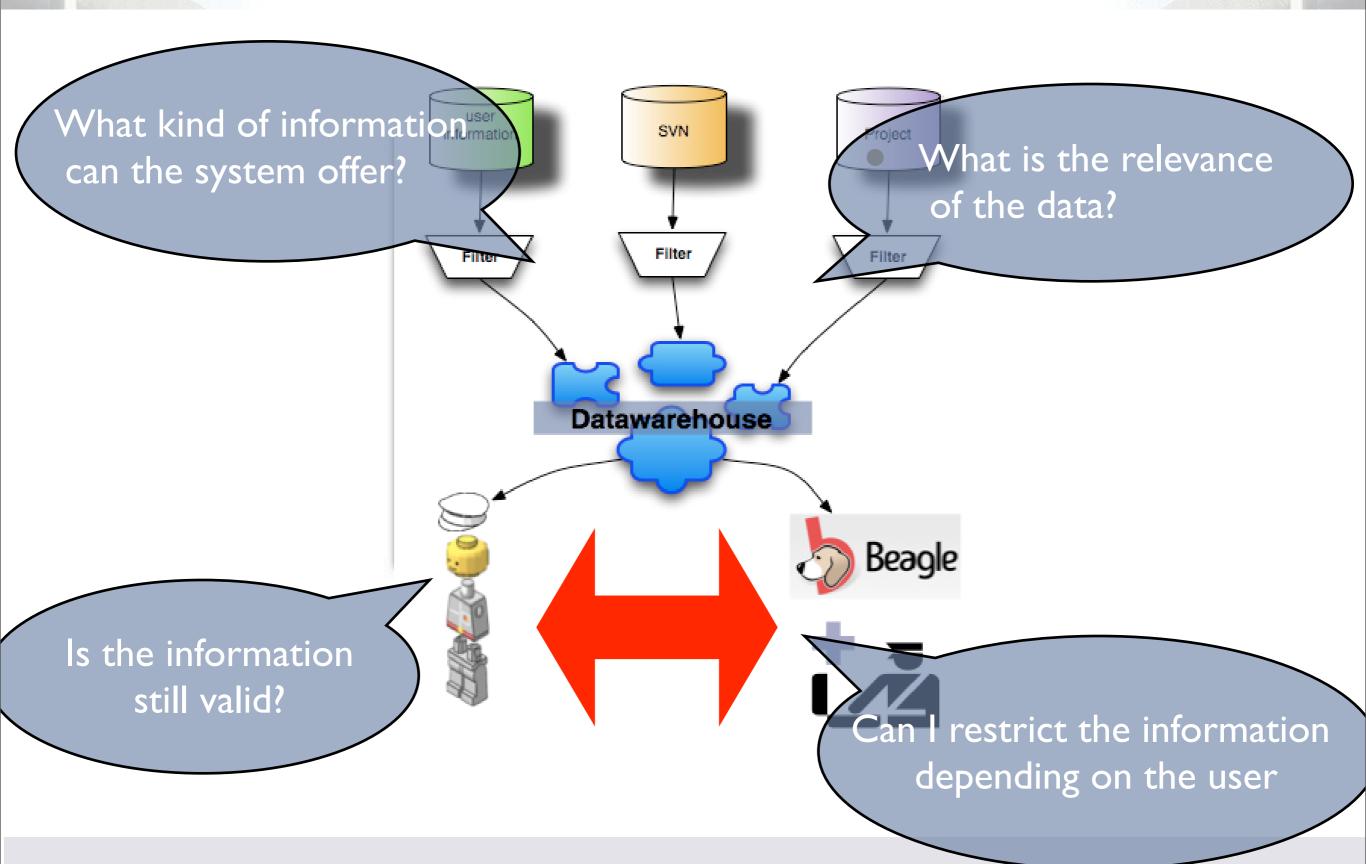




C/S-Tools	Ansprechpartner	Telefon	Ort	Status
Ant	Constantine Help	(HH)-21 86	(HH)-J335	online
Eclipse/Tools	Alfred E. Neumann Constantine Help	(HH)-37 53 (HH)-21 86	(HH)-J335 HH-J335	online online







Ref[I]









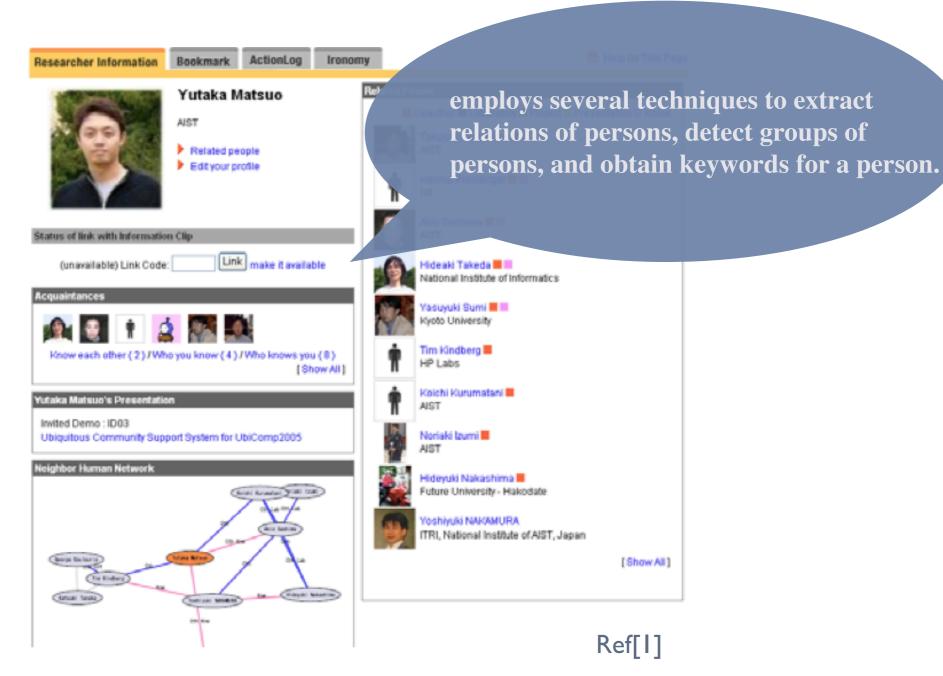




## What is Polyphonet







## What is Polyphonet







What is Polyphonet





- Class of Relation, more than 30 kinds of relationships
- Scalability,
   number of queries could behave like O(n^2)
- Co occurrence, personal meta-data useful for information retrieval & recommendations, compare person-to-word matrix



Algorithm for data mining



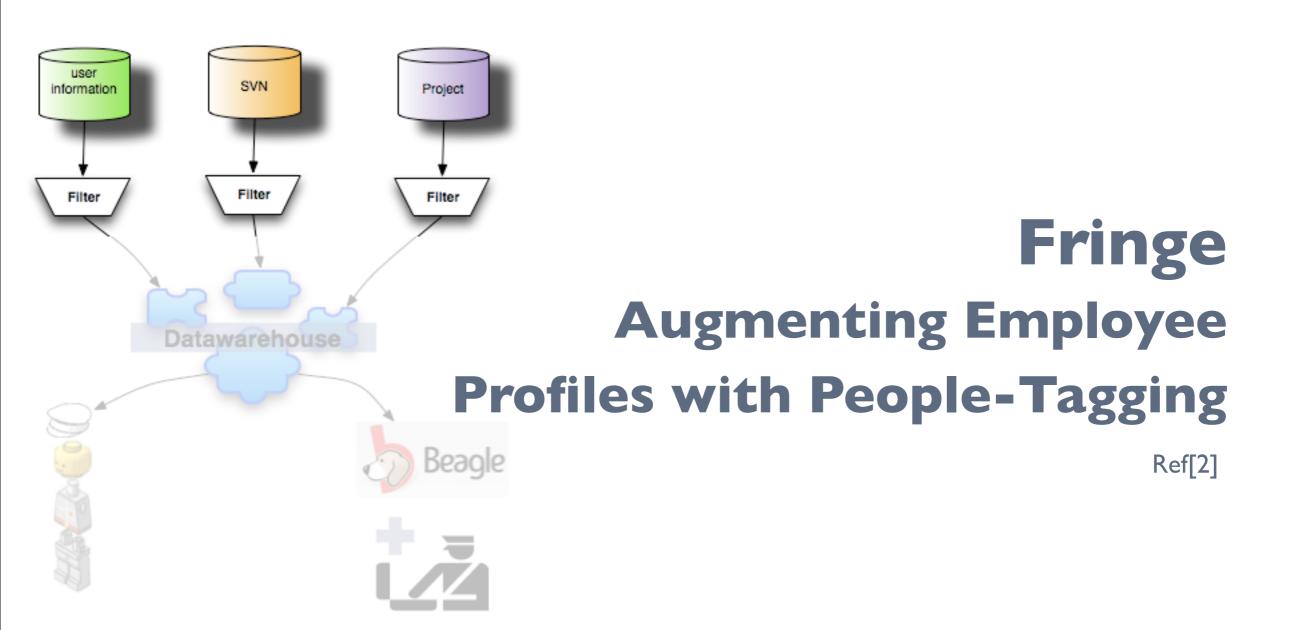


#### **Drawbacks**

- is even more than a proof of concept
- shows some implementation for extracting data
- depends only on web documents and relation between person
- does not include enough structural data







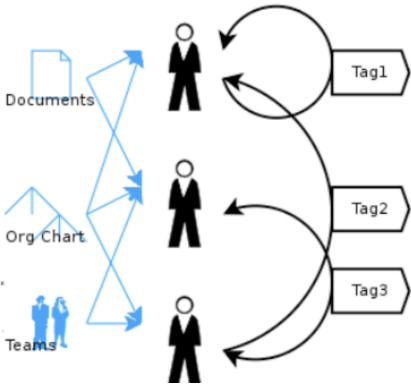




### What is Fringe?

 a system that enables users to tag other users with key words that are displayed on their profile

contributes distinctive information
 to the employees profile beyond the
 information available in the basic profile.





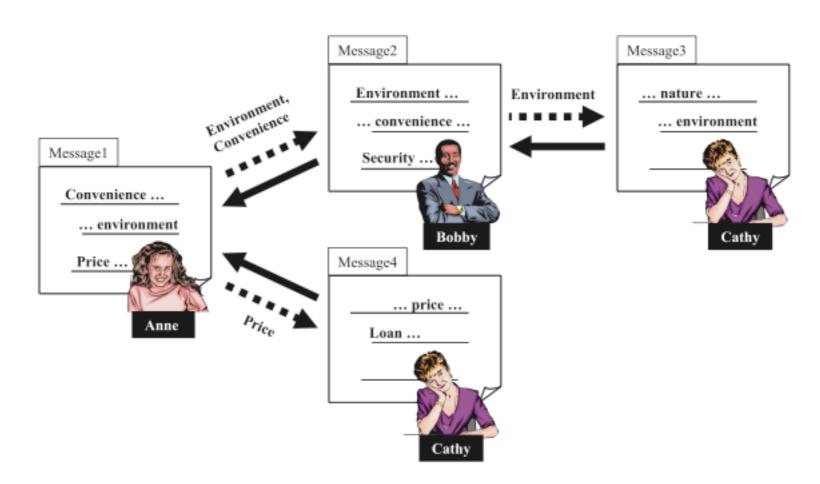


#### **Drawbacks**

- depends on all kind of information
- implements in all kind of tools are available
- is an easy Algorithm
- uses something similar to Google's Page Ranking, indicates the strength to a topic, depends on how often tags
- uses tags based on personal information, not verification



## Mining Directed Social Network from Message Board



Ref[3]

Ref[4]









### Two types of Information

- Relation between individual persons
- Any resource related to a person (web resources)

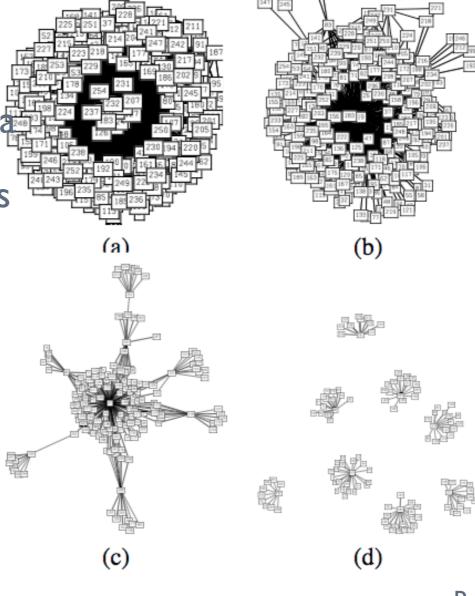




### Create a social network

 create (semi)automatically social network from a collection of web documents

- represent each person by
  - features
  - attributes



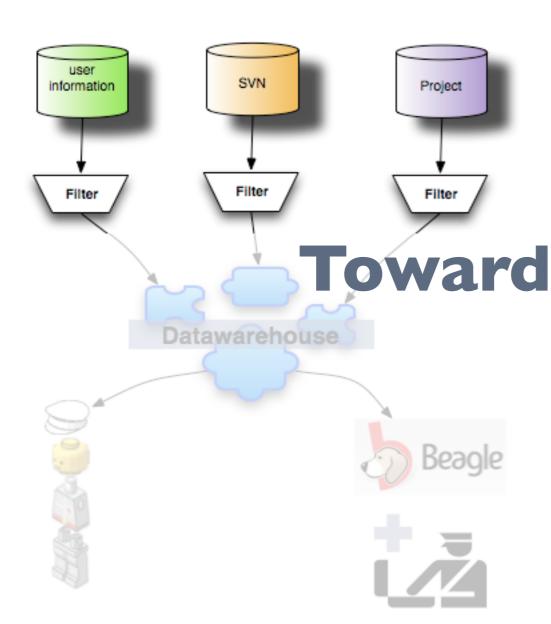




#### **Drawbacks**

- represent social networks by a adjacency matrix
- propose an aggregation vector instead of document similarity
- in the social network is a low precision
- Classify the Support Vector Machines (SVM)
- Information retrieval

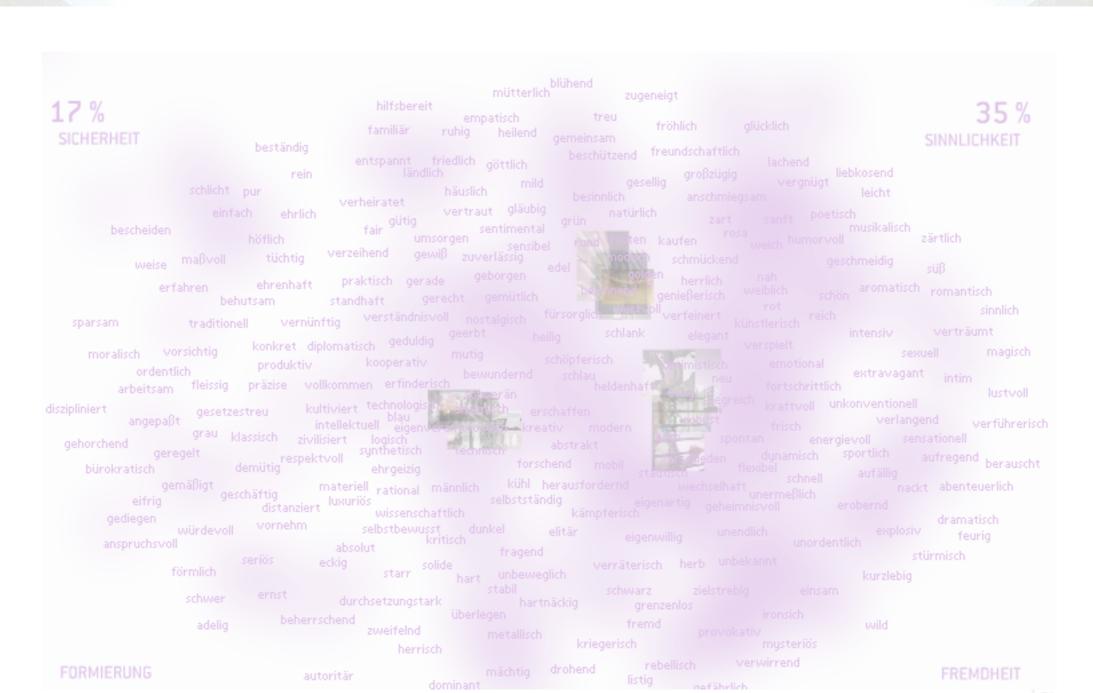






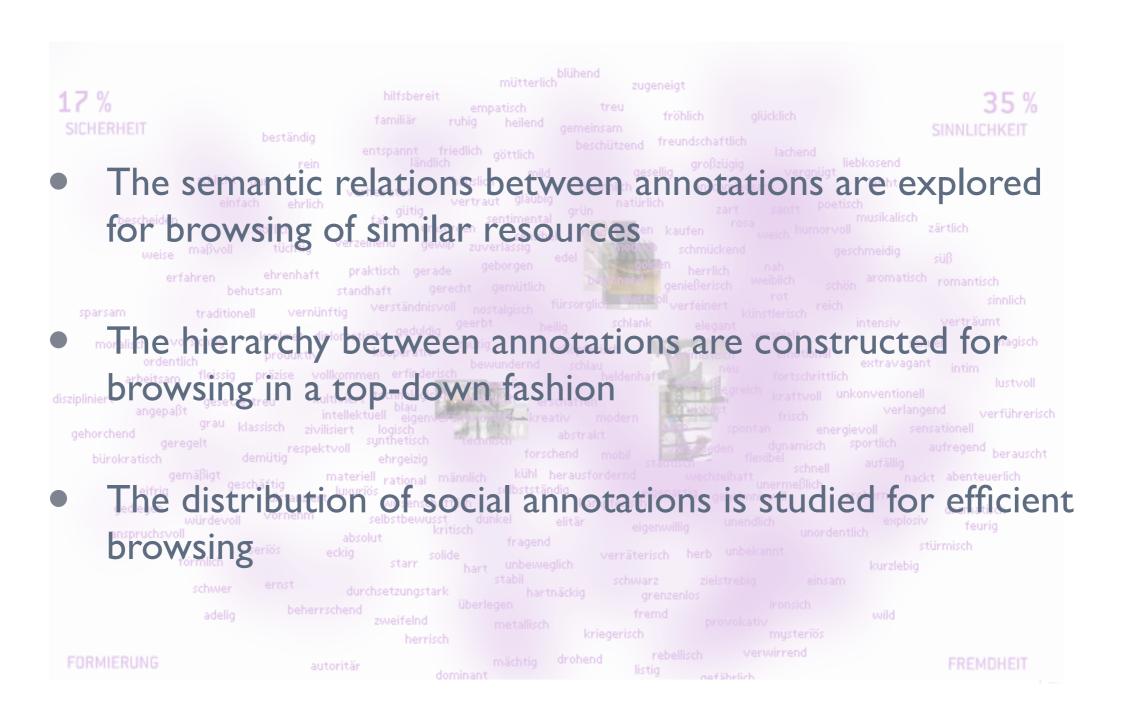
Ref[5]











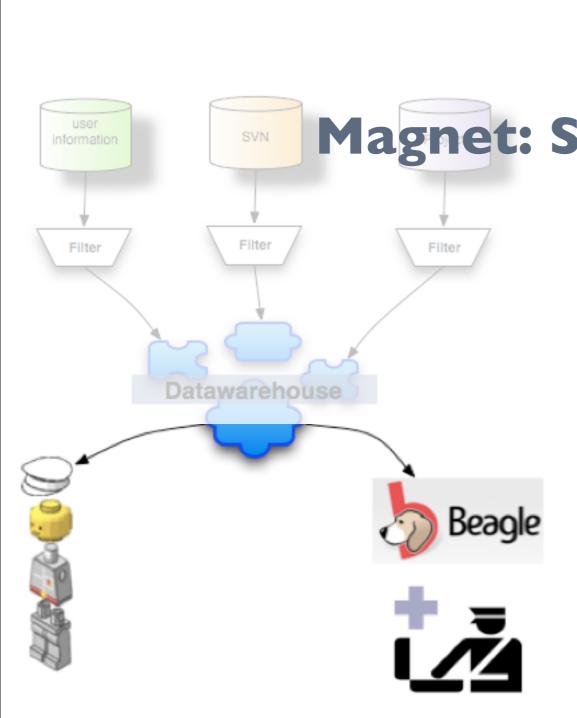




#### **Drawbacks**

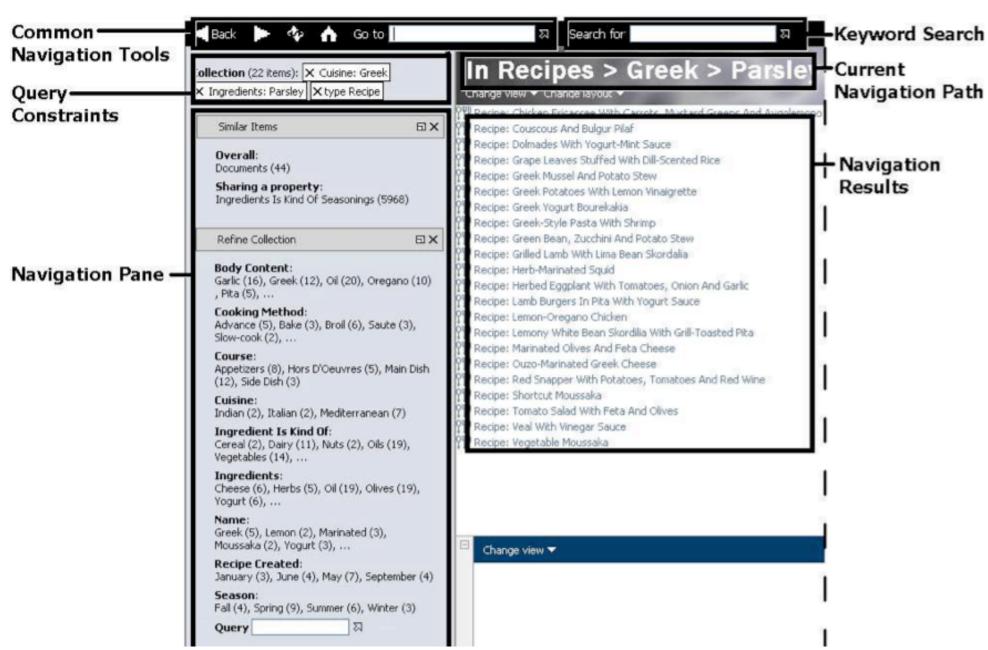
- describe some Cluster Analysis
- use Lucene API (like Beagle)
- time related browsing
- no implementation
- performance is unknown





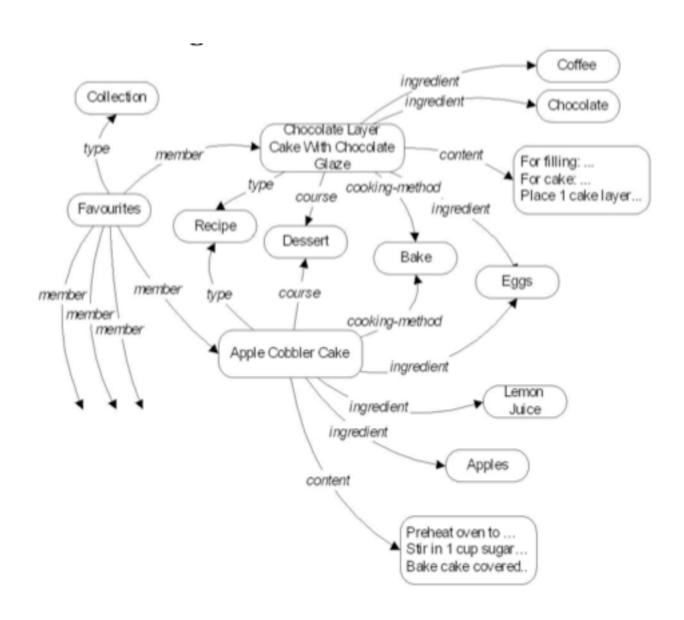








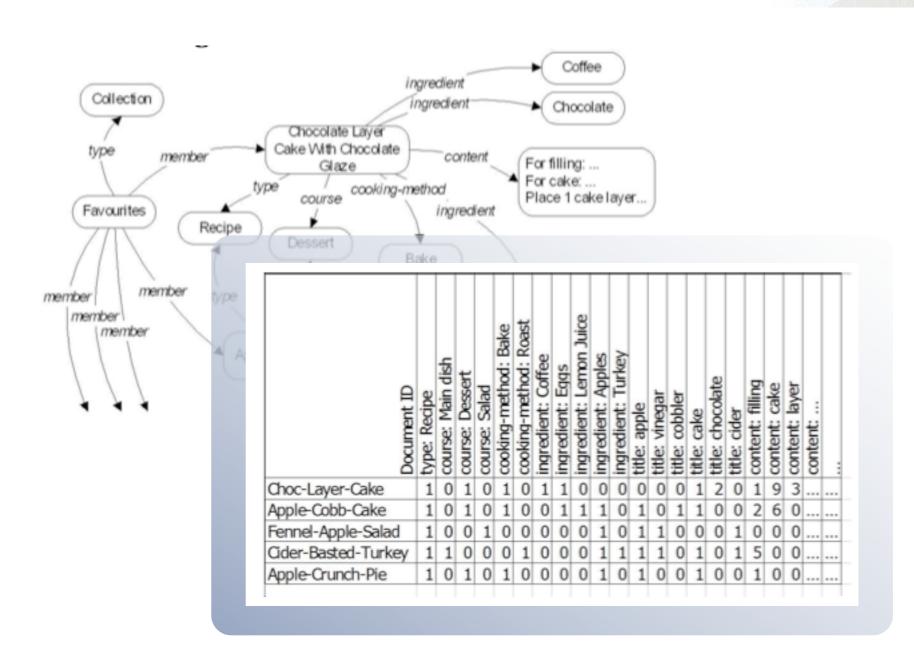




# Building a model







# Building a model

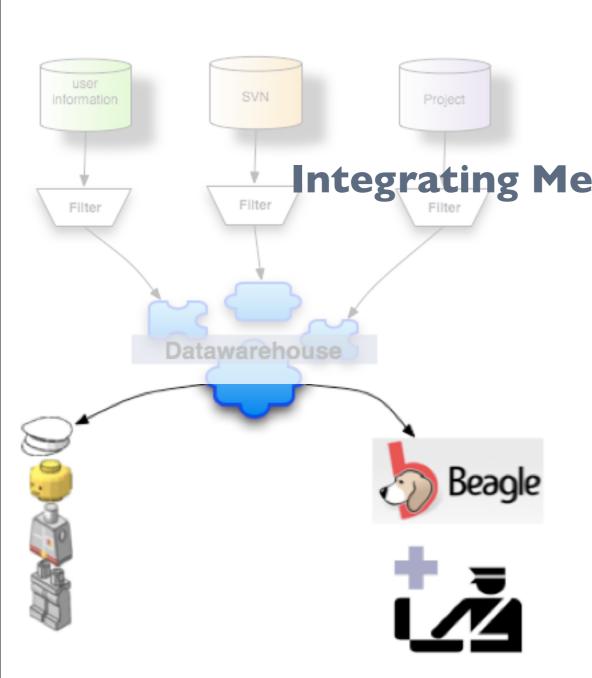


#### **Drawbacks**

- use RDF and Lucene
- build relations between different topics
- hard to work with negation (not in)
- no representative test with large scale data







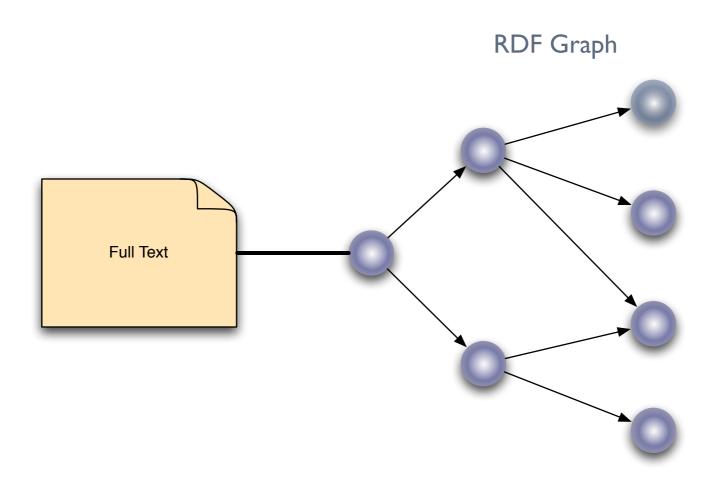
Keywords and RDF Fragments:
Integrating Meta-data and Full-Text Search in
Beagle++

Ref[7]





#### Document



Ref[7]

# Searchable Document, full text and reachable meta-data





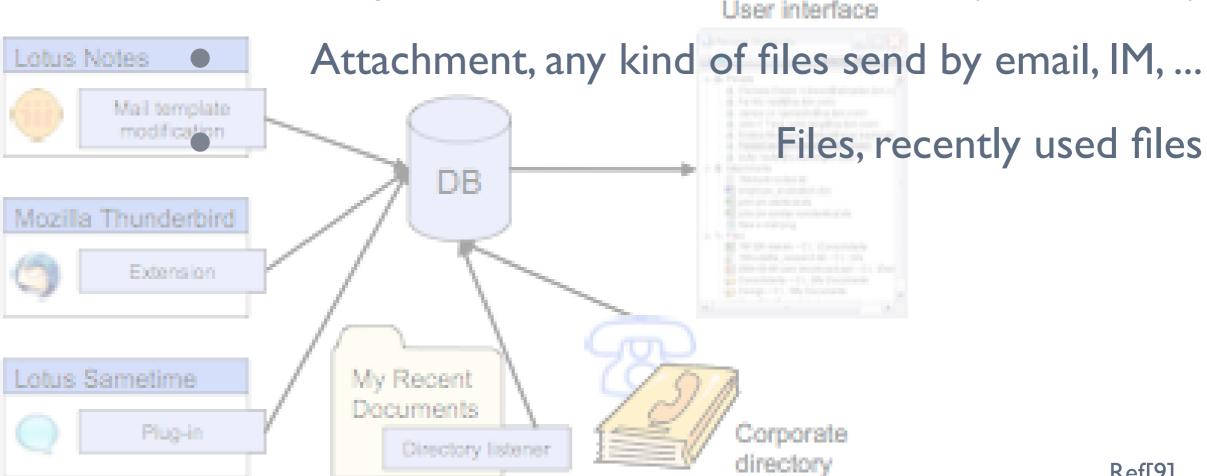
# Recent Shortcuts: Using Interactions to Support Shared Activities Ref[9]

- referencing to other people
- finding files to attach to email
- managing incoming email attachments
- managing varieties of shared files
- keeping track of other related work



### **Initial Prototype**

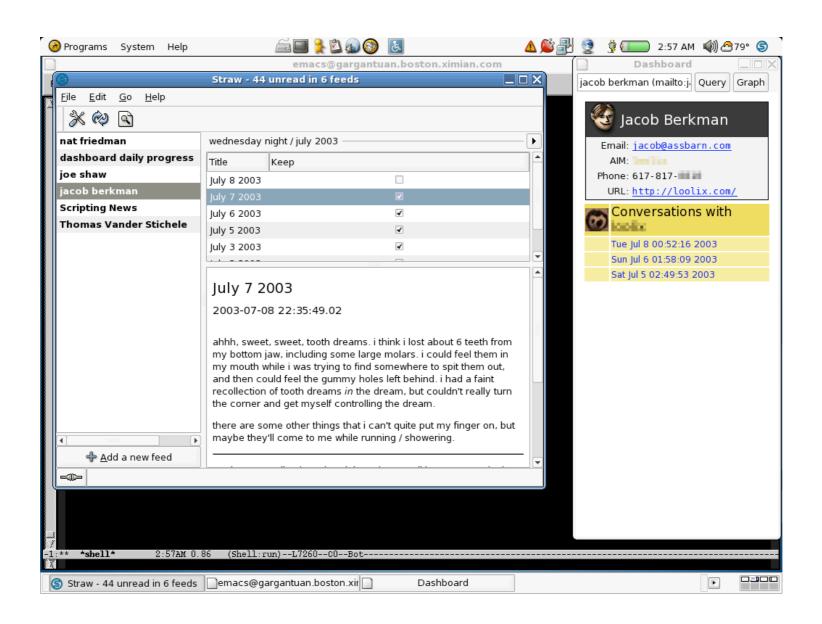
• People, interact with different tools (email, IM, ...)











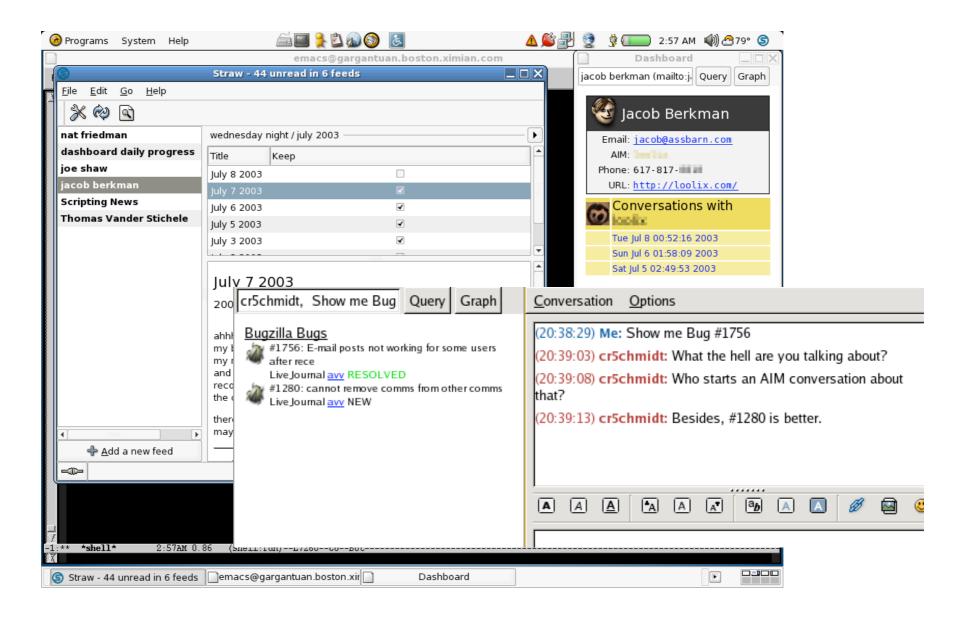
Why can't my computer automatically show me things that will help me with what I'm doing, instead of making me search around for them?

Ref[10]





#### **Dashboard**

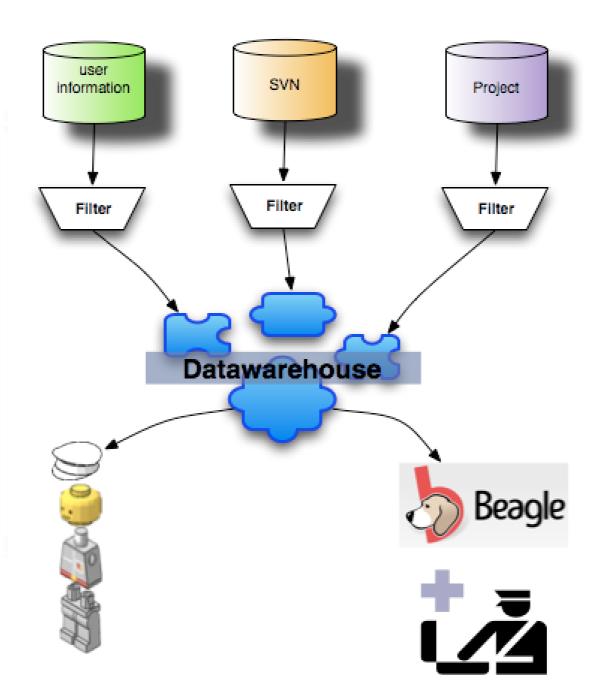


Why can't my computer automatically show me things that will help me with what I'm doing, instead of making me search around for them?

Ref[10]



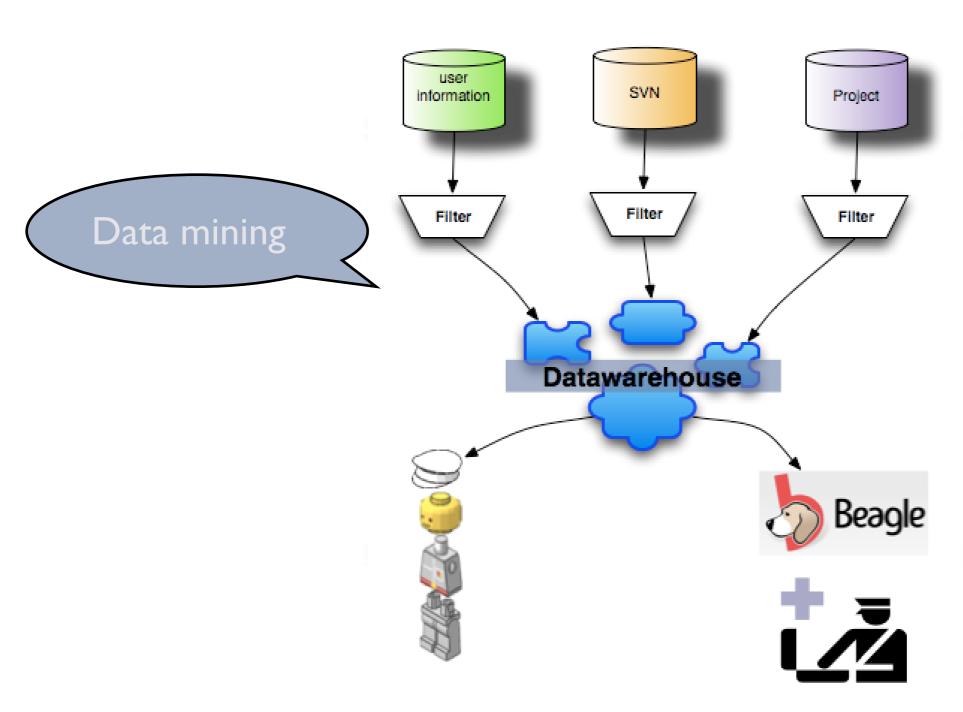




# Summary



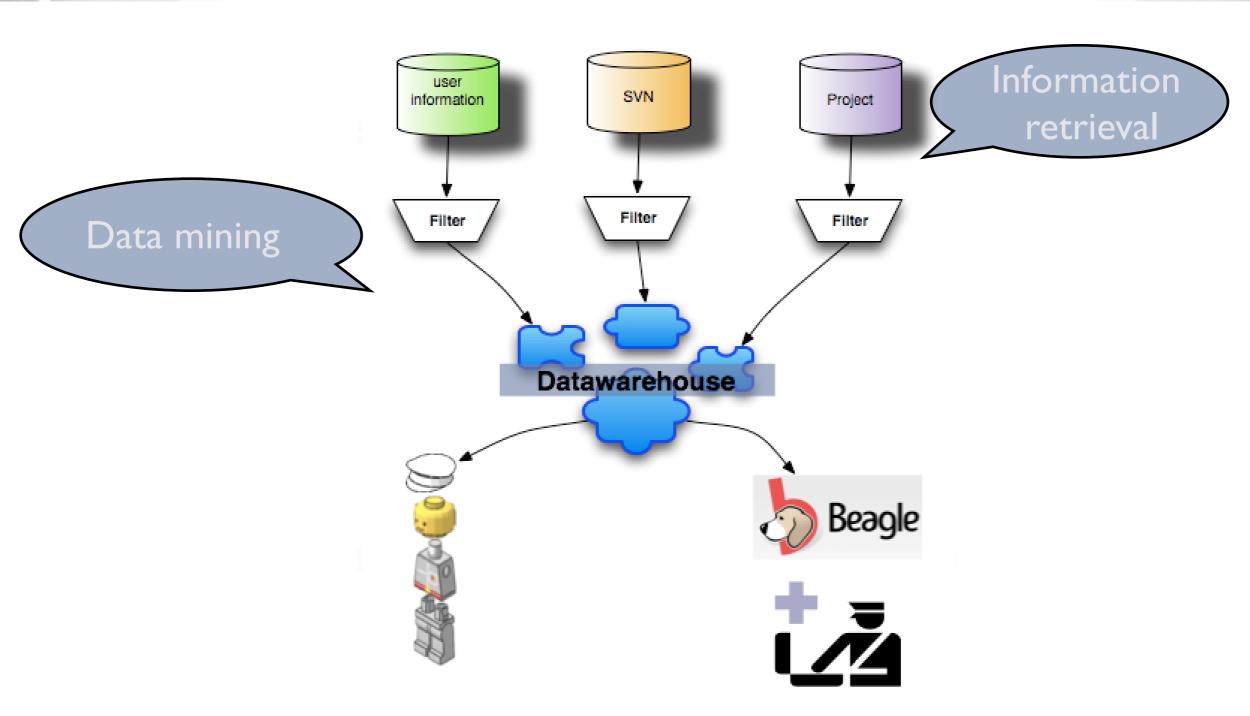




# Summary



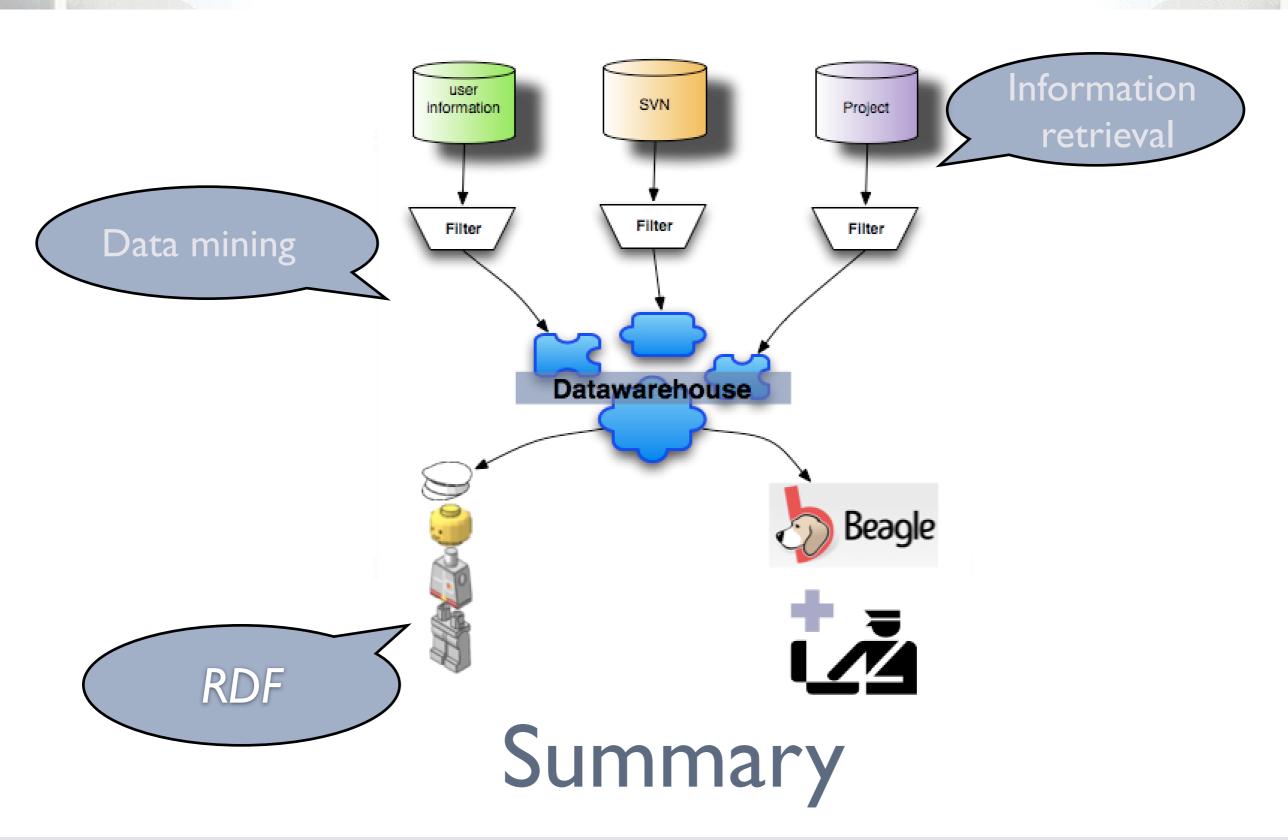




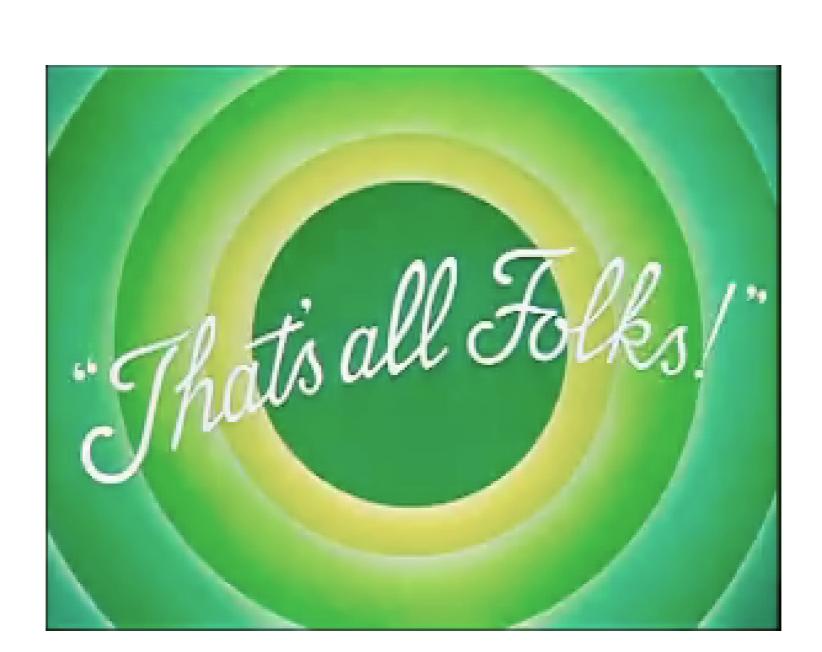
# Summary















- Matsuo, Y., Mori, J., Hamasaki, M., Ishida, K., Nishimura, T., Takeda, H., Hasida, K., and Ishizuka, M. 2006. POLYPHONET: an advanced social network extraction system from the web. In Proceedings of the 15th international Conference on World Wide Web (Edinburgh, Scotland, May 23 26, 2006). WWW '06. ACM, New York, NY, 397-406.
- Farrell, S., Lau, T., Nusser, S., Wilcox, E., and Muller, M. 2007. Socially augmenting employee profiles with people-tagging. In Proceedings of the 20th Annual ACM Symposium on User interface Software and Technology (Newport, Rhode Island, USA, October 07 10, 2007). UIST '07. ACM, New York, NY, 91-100.
- [3] Matsumura, N., Goldberg, D. E., and Llorà, X. 2005. Mining directed social network from message board. In Special interest Tracks and Posters of the 14th international Conference on World Wide Web (Chiba, Japan, May 10 14, 2005). WWW '05. ACM, New York, NY, 1092-1093.
- [4] Makrehchi, M. and Kamel, M. S. 2006. Learning Social Networks from Web Documents Using Support Vector Classifiers. In Proceedings of the 2006 IEEE/WIC/ACM international Conference on Web intelligence (December 18 22, 2006). Web Intelligence. IEEE Computer Society, Washington, DC, 88-94.





- Li, R., Bao, S., Yu, Y., Fei, B., and Su, Z. 2007. Towards effective browsing of large scale social annotations. In Proceedings of the 16th international Conference on World Wide Web (Banff, Alberta, Canada, May 08 12, 2007). WWW '07. ACM, New York, NY, 943-952.
- [6] Sinha, V. and Karger, D. R. 2005. Magnet: supporting navigation in semistructured data environments. In Proceedings of the 2005 ACM SIGMOD international Conference on Management of Data (Baltimore, Maryland, June 14 16, 2005). SIGMOD '05. ACM, New York, NY, 97-106.
- Tereza Iofciu, Christian Kohluetter, Raluca Paiu and Wolfgang Nejdl.Keywords and RDF Fragments: Integrating Metadata and Full-Text Search in Beagle++. Workshop on The Semantic Desktop Next Generation Personal Information Management and Collaboration Infrastructure at the International Semantic Web Conference 6 November 2005, Galway, Ireland





### Literature 3

[8]

Carmel, D., Maarek, Y. S., Mandelbrod, M., Mass, Y., and Soffer, A. 2003. Searching XML documents via XML fragments. In Proceedings of the 26th Annual international ACM SIGIR Conference on Research and Development in information Retrieval (Toronto, Canada, July 28 - August 01, 2003). SIGIR '03. ACM, New York, NY, 151-158.

- [9]
  - Tang, J. C., Lin, J., Pierce, J., Whittaker, S., and Drews, C. 2007. Recent shortcuts: using recent interactions to support shared activities. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (San Jose, California, USA, April 28 May 03, 2007). CHI '07. ACM, New York, NY, 1263-1272.
- [10]

Nat Friedman 2003. Dashboard. In Proceedings of the Linux Symposium Conference on the Linux kernel & major OS infrastructure and research projects, July 23rd-26th, 2003, Ottawa Canada, http://www.nat.org/dashboard/.