Deliver who I mean

Lars Maehlmann

Hamburg, 04.12.07
A Result

Suchseite – SIGNAL IDUNA info.net

Ihre Suchanfrage "ant eclipse" lieferte Ergebnisse in 2 Ansprechpartner, 2 Menüpunkten und 4 Dokumenten.

<table>
<thead>
<tr>
<th>C/S-Tools</th>
<th>Ansprechpartner</th>
<th>Telefon</th>
<th>Ort</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ant</td>
<td>Constantine Help</td>
<td>(HH)-21 86</td>
<td>(HH)-335 online</td>
<td></td>
</tr>
<tr>
<td>Eclipse/Tools</td>
<td>Alfred E. Neumann</td>
<td>(HH)-37 53</td>
<td>(HH)-335 online</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Constantine Help</td>
<td>(HH)-21 86</td>
<td>HH-335 online</td>
<td></td>
</tr>
</tbody>
</table>

Suchergebnisse 1 - 8 von 10

Menüpunkte

- Ressorts und LD/FDen > Ressort A > aeh-97400 > aems-97440 > Mitarbeiter
  Erstellt am 11.08.2005 / letzte Änderung 30.05.2006

- Ressorts und LD/FDen > Ressort B > aeh-97400 > aems-97440 > betreute Anwendungen
  Erstellt am 06.07.2005 / letzte Änderung 14.03.2007

Dokumente

- Eclipse
  Erstellt am 17.11.2006 / letzte Änderung 20.11.2006 von Constantine Help

- ant.htm
  Erstellt am 07.08.2002 / letzte Änderung 05.02.2005 von Constantine Help

- ADP_MB_Sektor14_17_PWNW.pdf
  Erstellt am 17.08.2001 / letzte Änderung 02.01.2006 von Alfred E. Neumann
A Result

Use case
## A Phone Number

<table>
<thead>
<tr>
<th>C/S-Tools</th>
<th>Ansprechpartner</th>
<th>Telefon</th>
<th>Ort</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ant</td>
<td>Constantine Help</td>
<td>(HH)-21 86</td>
<td>(HH)-J335</td>
<td>online</td>
</tr>
<tr>
<td>Eclipse/Tools</td>
<td>Alfred E. Neumann</td>
<td>(HH)-37 53</td>
<td>(HH)-J335</td>
<td>online</td>
</tr>
<tr>
<td></td>
<td>Constantine Help</td>
<td>(HH)-21 86</td>
<td>HH-J335</td>
<td>online</td>
</tr>
</tbody>
</table>
What kind of information can the system offer?

What is the relevance of the data?

Is the information still valid?

Can I restrict the information depending on the user?
Polyphonet
An Advanced Social Network Extraction System from the Web
Ref[1]
What is Polyphonet
What is Polyphotonet

employs several techniques to extract relations of persons, detect groups of persons, and obtain keywords for a person.
What is Polyphonnet

Search engines, especially Google, are used to measure co-occurrence of information and obtain Web documents networks.

Polyphonet employs several techniques to extract relations of persons, detect groups of persons, and obtain keywords for a person.
• 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
Fringe
Augmenting Employee Profiles with People-Tagging

Ref[2]
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]
Learning Social Networks from Web Documents Using Support Vector Classifiers

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 a social network from a collection of web documents

- represent each person by
  - features
  - attributes

Ref[4]
Drawbacks

- represent social networks by an 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
Towards effective browsing of large scale social annotations

Ref[5]
Towards effective browsing
• The semantic relations between annotations are explored for browsing of similar resources.

• The hierarchy between annotations are constructed for browsing in a top-down fashion.

• The distribution of social annotations is studied for efficient browsing.
Drawbacks

- describe some Cluster Analysis
- use Lucene API (like Beagle)
- time related browsing
- no implementation
- performance is unknown
Magnet: Supporting Navigation in Semistructured Data Environments

Ref[6]
Building a model
Building a model

Ref[6]
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]
Searchable Document, full text and reachable meta-data
Recent Shortcuts: Using Interactions to Support Shared Activities

- 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, ...)
- Attachment, any kind of files sent by email, IM, ...
- Files, recently used files
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?
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?
Summary
Data mining

Summary
Summary

Data mining

Information retrieval

Summary
Data mining

Information retrieval

RDF

Summary
“That's all Folks!”


