

Intelligente Systeme im World Wide Web

Anwendungen

Dr. York Sure

Institut AIFB, Universität Karlsruhe

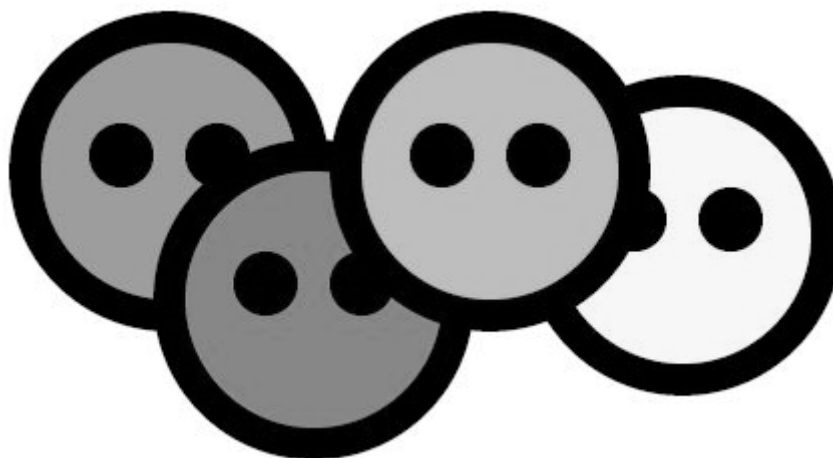
Inhalt

- RDF – Friend of a Friend
- RDF Schema – Bibster

- OWL – Web Service Matching
- F-Logic – Halo Chemistry

the foaf project (RDF)
see <http://www.foaf-project.org/>

The *Friend of a Friend* (FOAF) project is about creating a Web of machine-readable homepages describing people, the links between them and the things they create and do.



<http://www.ldodds.com/foaf/foaf-a-matic.html>

- FOAF-A-MATIC
 - <http://www.ldodds.com/foaf/foaf-a-matic.html>
 - Einfaches Javascript
 - Generieren von FOAF-Files
- FOAFNAUT
 - <http://foafnaut.org/>
- FoaF Explorer, usw. ...
 - <http://xml.mfd-consult.dk/foaf/explorer/>

FOAF File example

```
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xmlns:foaf="http://xmlns.com/foaf/0.1/"
  xmlns:admin="http://webns.net/mvcb/">
<foaf:PersonalProfileDocument rdf:about="">
  <foaf:maker rdf:nodeID="me"/>
  <foaf:primaryTopic rdf:nodeID="me"/>
  <admin:generatorAgent
    rdf:resource="http://www.ldodds.com/foaf/foaf-a-matic"/>
  <admin:errorReportsTo rdf:resource="mailto:leigh@ldodds.com"/>
</foaf:PersonalProfileDocument>
```

```
<foaf:Person rdf:nodeID="me">  
<foaf:name>York Sure</foaf:name>  
<foaf:title>Dr.</foaf:title>  
<foaf:givenname>York</foaf:givenname>  
<foaf:family_name>Sure</foaf:family_name>  
<foaf:mbox_sha1sum>c013b5c247e106e7092b1fce764  
502052646e30d</foaf:mbox_sha1sum>  
<foaf:homepage rdf:resource="http://www.aifb.uni-  
karlsruhe.de/WBS/ysu"/>  
<foaf:depiction rdf:resource="http://www.aifb.uni-  
karlsruhe.de/WBS/ysu/images/me.jpg"/>
```

```
<foaf:phone rdf:resource="tel:+49-721-608-6592"/>
<foaf:workplaceHomepage rdf:resource="http://www.aifb.uni-
karlsruhe.de/WBS"/>
<foaf:workInfoHomepage rdf:resource="http://www.aifb.uni-
karlsruhe.de/WBS/ysu"/>
<foaf:knows>
<foaf:Person>
<foaf:name>Pascal Hitzler</foaf:name>
<foaf:mbox_sha1sum>ab5b33f1535f1e3460024c5cd25bd72bffb53
dbf</foaf:mbox_sha1sum></foaf:Person></foaf:knows></foaf:P
erson>
</rdf:RDF>
```

foafnaut - Microsoft Internet Explorer

Datei Bearbeiten Ansicht Favoriten Extras ?

Adresse <http://www.foafnaut.org/2003/snap1/foafnaut.svg?email=danbri@w3.org> Wechseln zu Links >>

Google Web-Suche 0 blockiert Optionen

The main content is a network graph titled "foafnaut". At the center is a node labeled "DANBRI". Radiating from this central node are numerous dashed lines connecting to other nodes. These nodes are arranged in two vertical columns. The left column contains names: DAJOBE, MICHAEL, JAN, EM, JOE, MARTIN, EDD, JIBBERJIM, IAN, DAMEY, and LIBBY. The right column contains names: NMG, CHARLES, NICOLE, CHRIS, HEX, HITHERTO, NADIA, EIKEON, BWM, BRIAN, RALPH, and JHENDLER. Each node is represented by a small icon with a plus or minus sign. Three pop-up windows are visible:

- DANBRI**: Contains a portrait of a man.
- DANBRI**: Contains the following text:

DAN BRICKLEY
 danbri@w3.org
 daniel.brickley@bristol.ac.uk
 danbri@rdfweb.org

Knows: 16
 Codepictions: 25
 Known by: 24
 homepage
- DANBRI/JHENDLER**: Contains a group photograph of several people.

Internet

Bibster (RDF + RDFS)
Sharing Bibliographic Metadata in a P2P Network
<http://bibster.semanticweb.org>

Agenda

1. Introduction:
The Bibliographic Scenario
2. The Bibster System
System Architecture
Semantic Methods in Bibster
3. Evaluation

1. Introduction: Scenario

- **Scenario:** Sharing of bibliographic metadata in a Peer-to-Peer network
 - Bibliographic metadata is created and maintained in a **decentralized** manner, centralized solution not applicable
 - Researchers are willing to share their data
 - Use of semantics is crucial in this setting
- The Bibster system allows to:
 - **Easily share** bibliographic data
 - Save work in finding this data
 - Avoid re-typing this data by hand



2. Bibster Screenshot

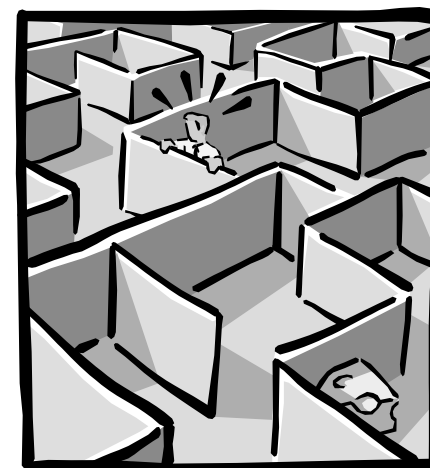
Query Scope

Semantic Search

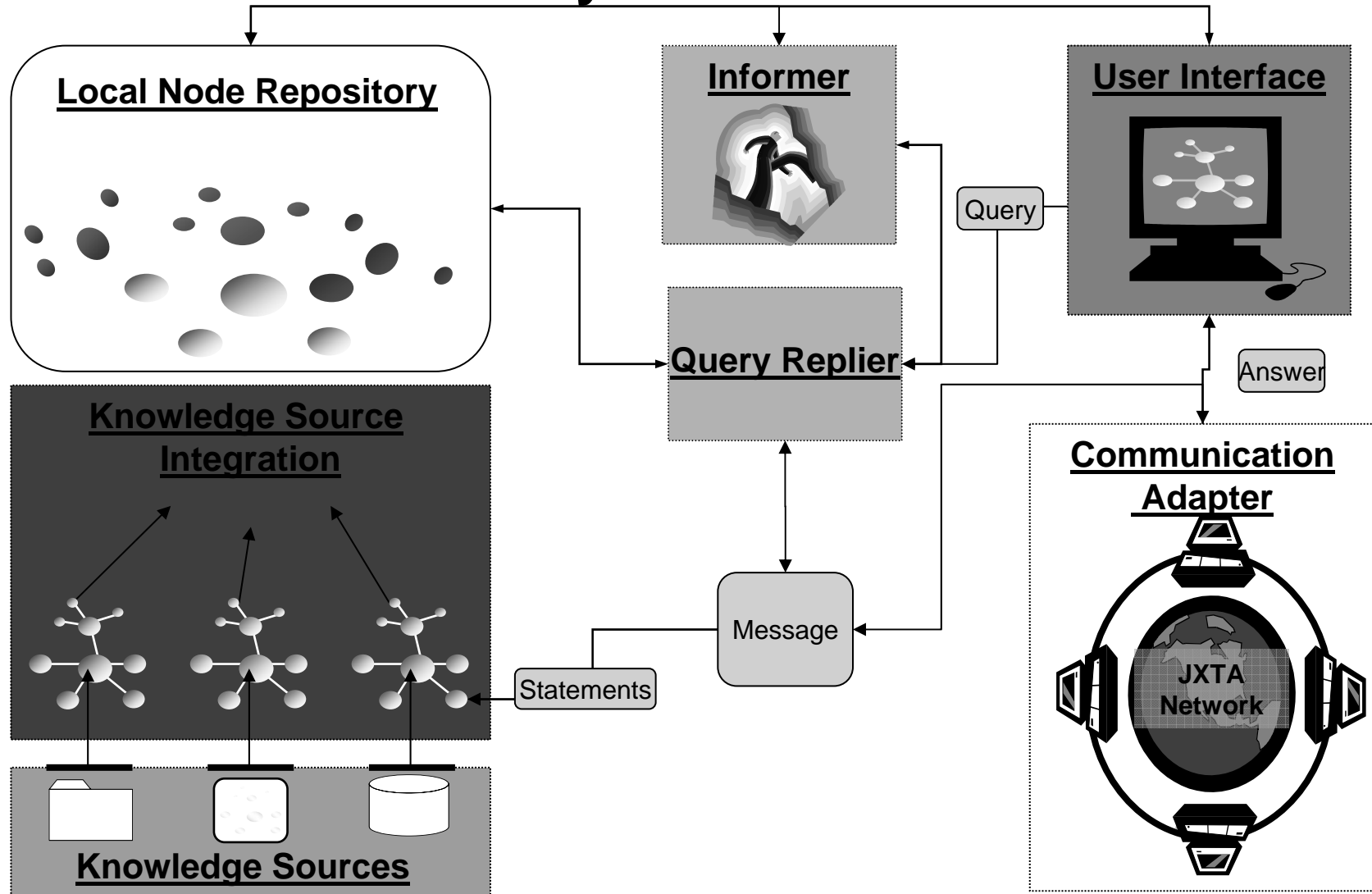
The screenshot shows the Bibster web interface. At the top, there is a table with columns: 'dfs:label', 'swrc:author', 'swrc:year', 'swrc:pages', and 'rdf:type'. The first row contains the text 'The Capabilities of Relational Database Management Systems', 'E. F. Codd', '1981', and 'Article'. Below this table, there are three radio buttons for 'Scope': 'Local Peers', 'All Peers' (which is selected), and 'Selected Peers'. A 'Change...' button is next to the 'Selected Peers' option. Below the scope options is the 'Search Details' section, which includes a list of publication types: 'Article', 'Book', 'Booklet', 'InBook', 'InCollection', and 'InProceedings'. Below this list are several input fields for search criteria: 'title', 'abstract', 'keywords', 'year', 'author' (with 'Codd' entered), 'journal', and 'url'. To the right of the search details is a large area labeled 'Query Results' with an upward-pointing arrow. Below the 'Query Results' area is a section labeled 'Integration and Export of Query Results' with a downward-pointing arrow. At the bottom of the screenshot, there is a 'BIB1 ex view' section showing a detailed query result for an article. The result is a list of key-value pairs: '@Article{1952311520', 'journal='IBM Research Report, San Jose, California'', 'isAbout='Data, Systems, Management, Database_Management'', 'key='DBLP:persons/Codd81'', 'volume='RJ3132'', 'ee='db/labs/ibm/RJ3132.html'', 'author='E. F. Codd'', 'label='The Capabilities of Relational Database Management Systems'', 'year='1981'', 'title='The Capabilities of Relational Database Management Systems'', and 'bibsource='DBLP, http://dblp.uni-trier.de''.

2. Semantic Methods in Bibster

- Semantic representation and querying of **metadata**
 - Extraction and classification from e.g. BibTeX files
 - **Semantic Web Research Community Ontology** and **ACM Topic hierarchy** as light-weight ontologies
- **Peer selection** using semantic topologies
 - Scalability requires intelligent query routing
 - Semantic descriptions of peers' expertise to build semantic topologies as basis for peer selection
- Semantic **duplicate detection**
 - Highly redundant and inconsistent representation of bibliographic metadata
 - Semantic similarity measures to detect duplicates

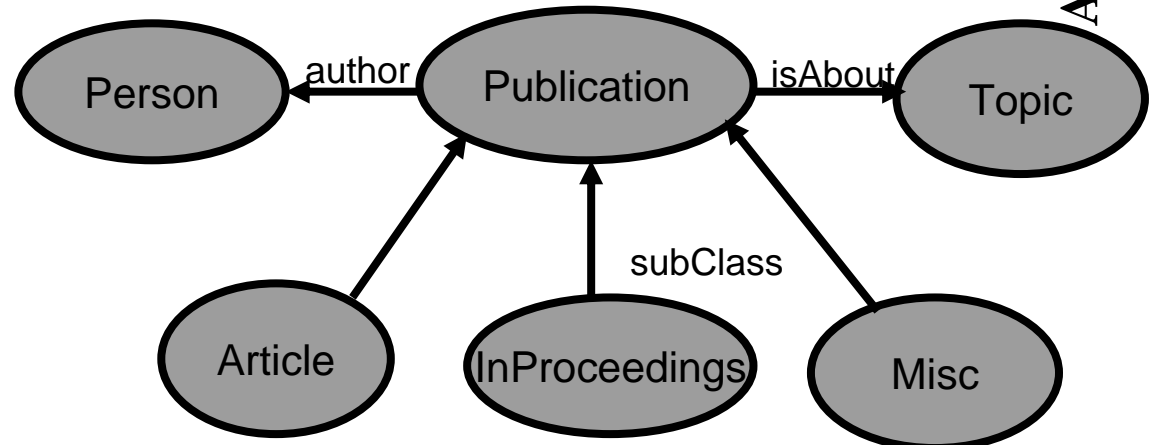


2. Bibster / SWAP System Architecture



Semantic Representation of Metadata

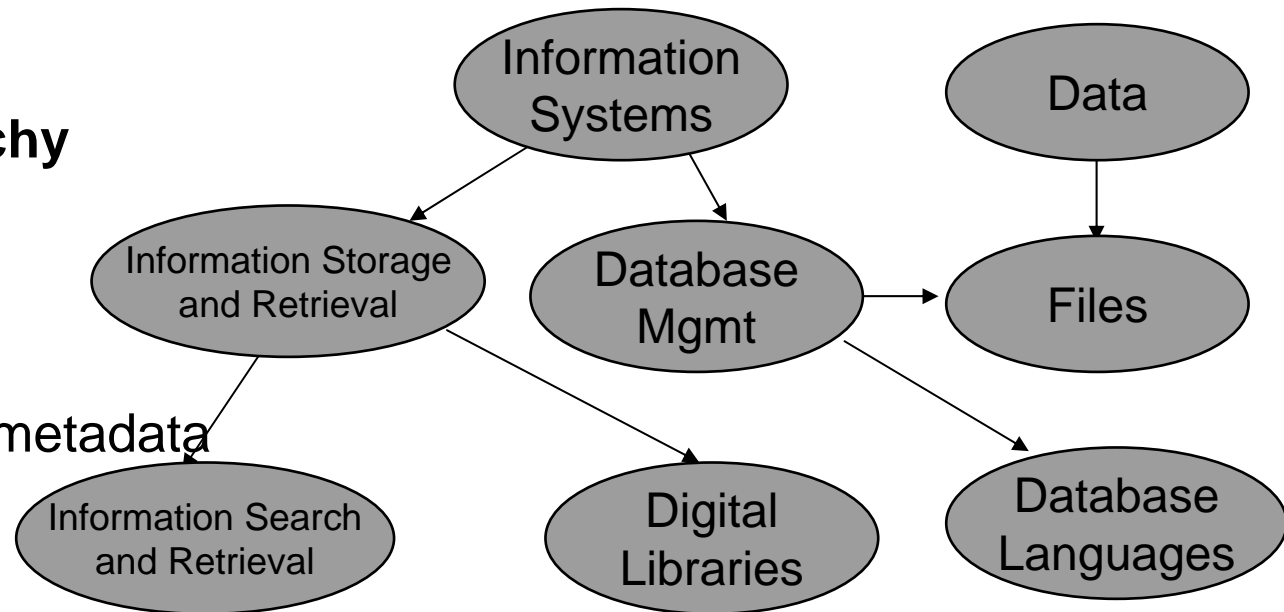
SWRC as base ontology:
models domain of research
community



ACM Topic hierarchy

1287 topics from
Computer Science
Domain

for classification of metadata
(linked by
hasSubtopic)

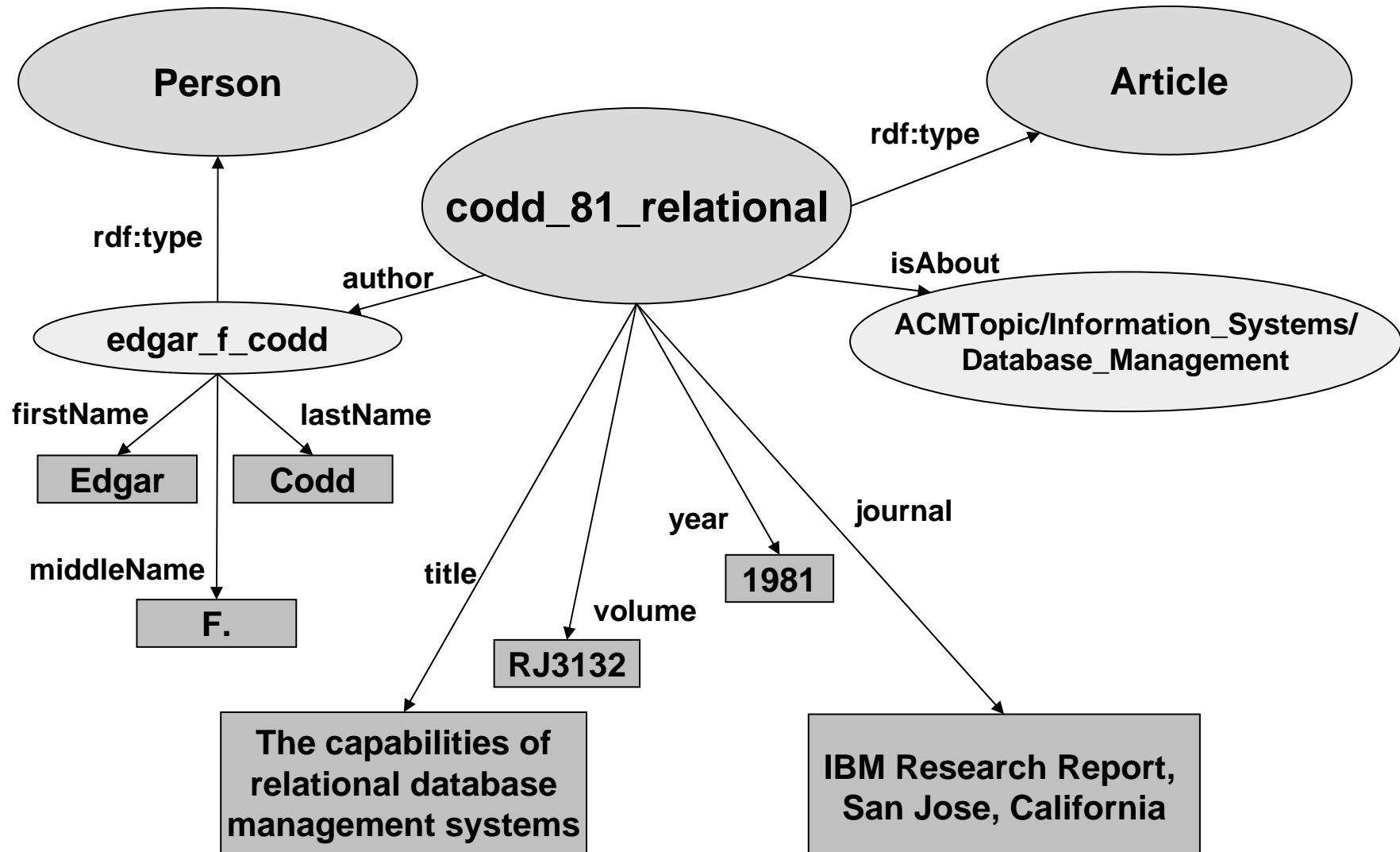


subTopic seeAlso

Sample BibTeX Entry

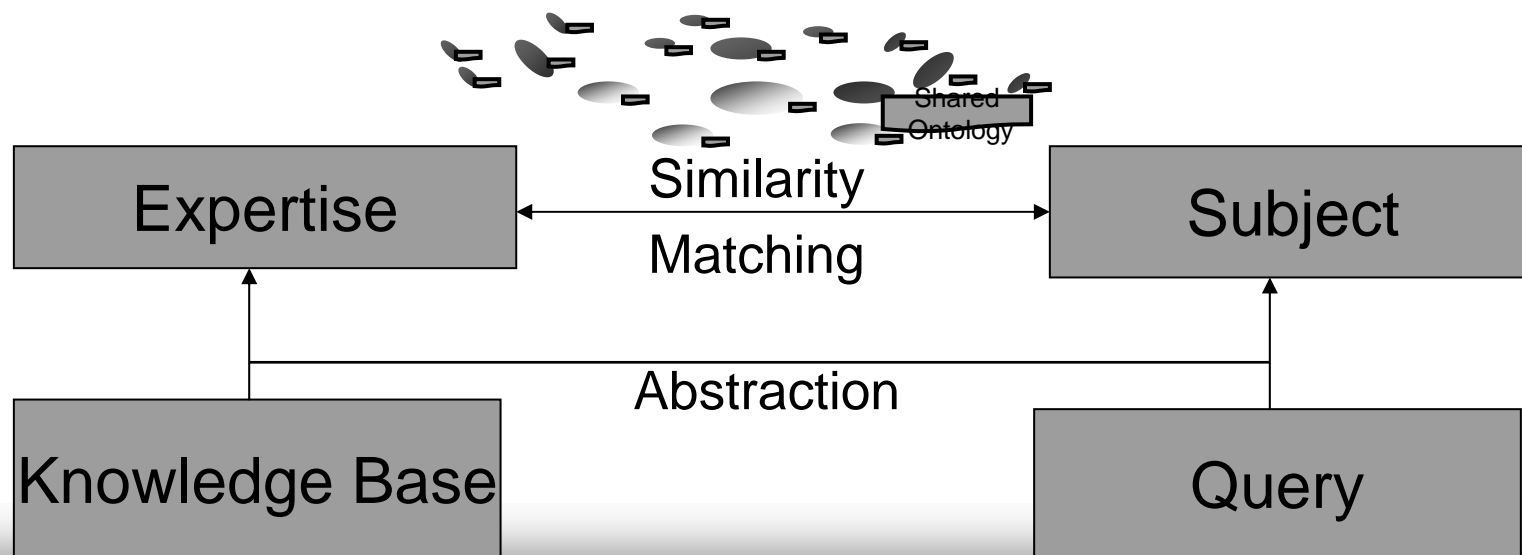
```
@ARTICLE{codd81relational,  
  author = {Edgar F. Codd},  
  title = {The capabilities of relational database  
    management systems},  
  journal = {IBM Research Report, San Jose, California},  
  volume = {RJ3132},  
  year = {1981}  
}
```

Sample Entry

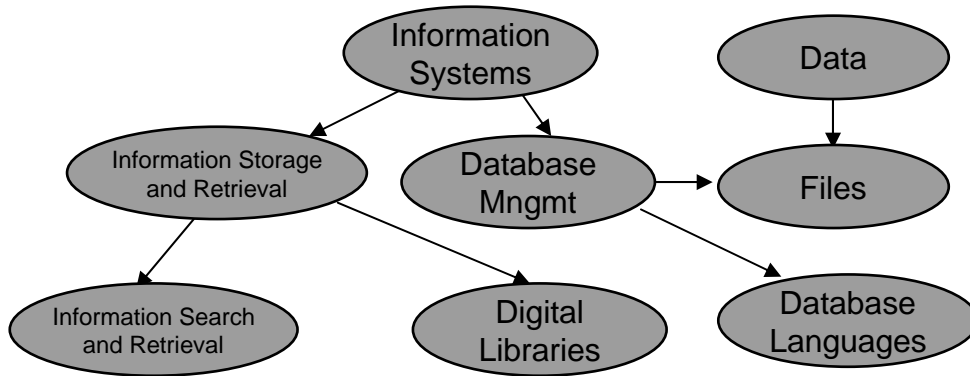


Expertise-Based Peer Selection

- **Advertisements** to promote semantic descriptions of expertise in the network
- **Semantic topology:** formed by the knowledge about other peers' expertise, peers with similar expertise are clustered
- **Peer Selection:** ranks peers according to similarity between their expertise and query subject

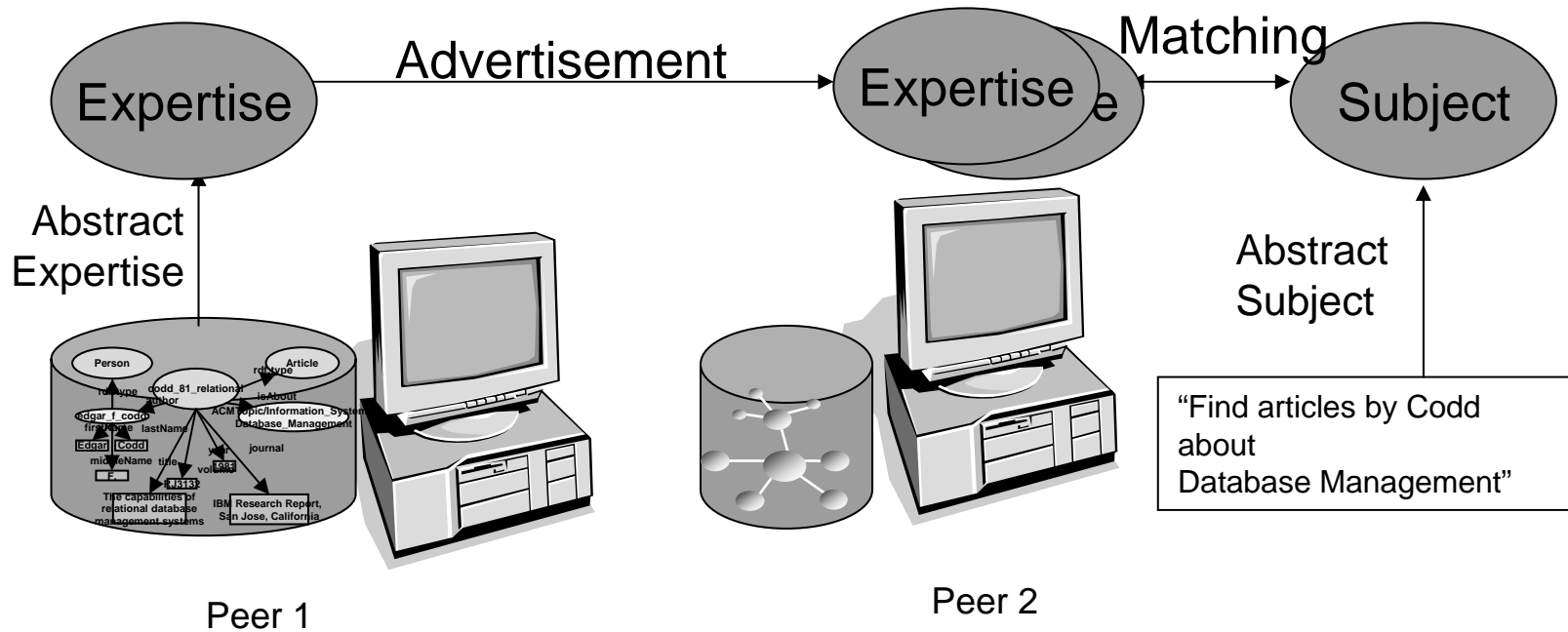


Expertise-Based Peer Selection

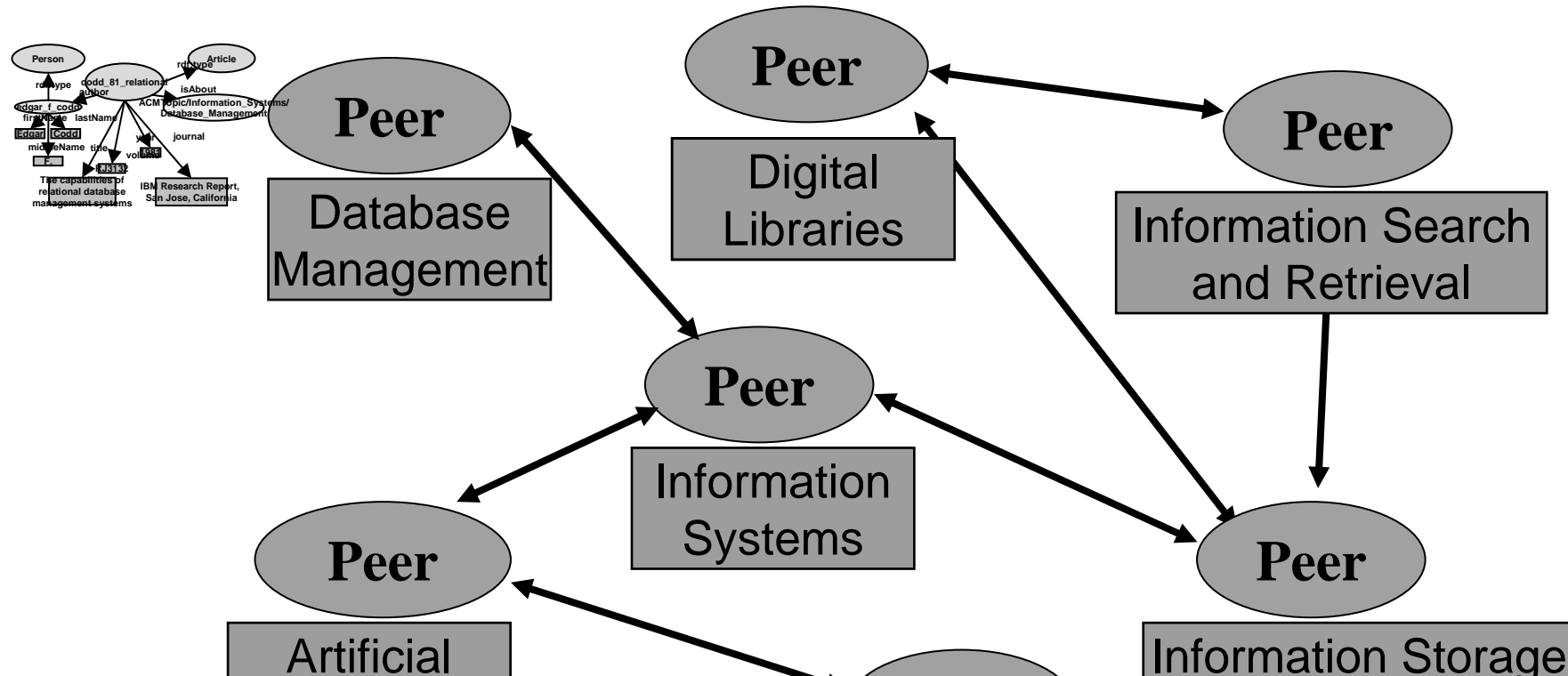


$$sim_{Topic}(t_1, t_2) = e^{-\alpha l} \cdot \frac{e^{\beta h} - e^{-\beta h}}{e^{\beta h} + e^{-\beta h}}$$

Similarity Function



Semantic Topologies

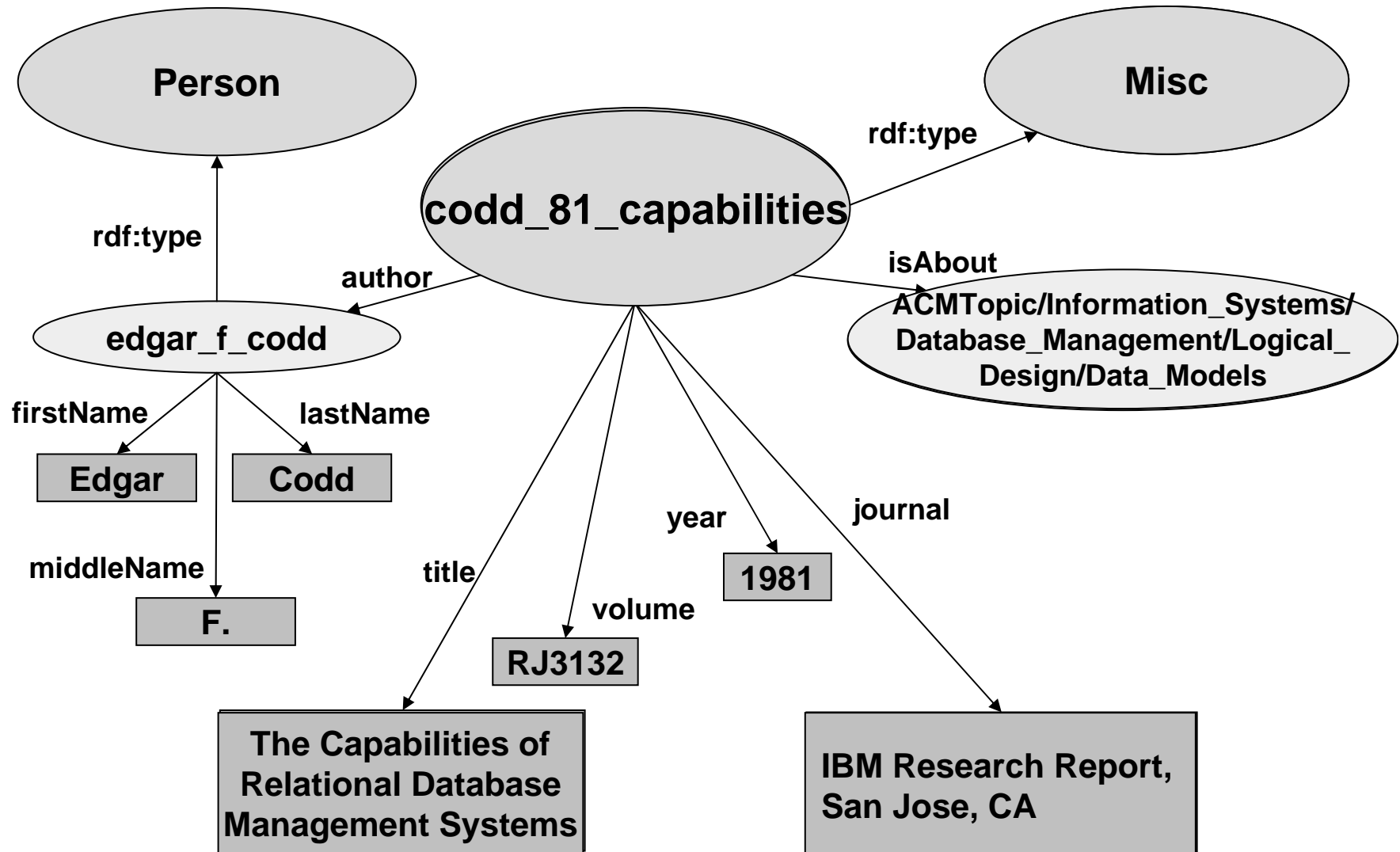


„Peer Selection in Peer-to-Peer Networks with Semantic Topologies“ (Haase, Siebes, van Harmelen)
 International Conference on Semantics for a Networked World, Paris, June 2004

Database Management”

ROBOTICS

Duplicate Entry



Semantic Duplicate Detection

- Individual similarity functions
 - Syntactic: Levenshtein similarity for titles etc.
 - Graph Structure: Sequence of authors
 - Ontological Structure: ACM Topic hierarchy
 - Domain Knowledge: e.g. Type Misc often corresponds to Unknown
- Aggregated similarity function:
 - e.g. weighted average
- Duplicates: Entries with similarity above a specified threshold
- Clusters of duplicate entries
- Merging of entries based on heuristics

3. Evaluation

- We arranged case studies with Bibster to
 - evaluated the scalability, functionality, and the performance of a semantics based Peer-to-Peer system
 - validated results from simulation experiments
- How did we evaluate?
 - User queries, routing information etc. were logged for evaluation purposes
 - After the case studies we provided a user questionnaire to evaluate the usability of our system
- Some findings:
 - Personalized recommendations outperform random selections
 - For more, see <http://bibster.semanticweb.org/>

4. Conclusion

- Fully implemented semantic P2P system for the bibliographic domain
- Exploitation ontologies in all steps:
 - Importing and representing the data
 - Querying the data
 - Routing requests
 - Integrating heterogeneous results
- Evaluation using simulation experiments and real-life case study