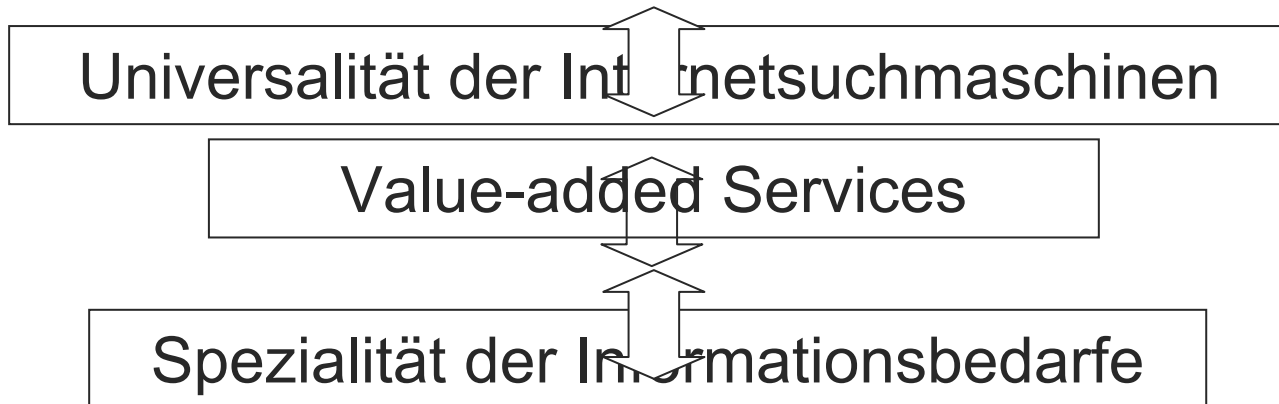


Metasuchmaschine für eine spezielle Anwendungsdomäne

Pavel Braslavski
06.02.2004

[Zielsetzung]



Möglichkeiten der Spezialisierung

- Anfrageoperationen
(Erweiterung, Modifikation, Übersetzung usw.)
- Bildung der Rangliste
(allerdings kein „globales Wissen“ vorhanden)

[3 in 1]

- Metasuche
- Domänenspezifisches Wissen = Thesaurus
- Visualisierung (grafisches Suchinterface)

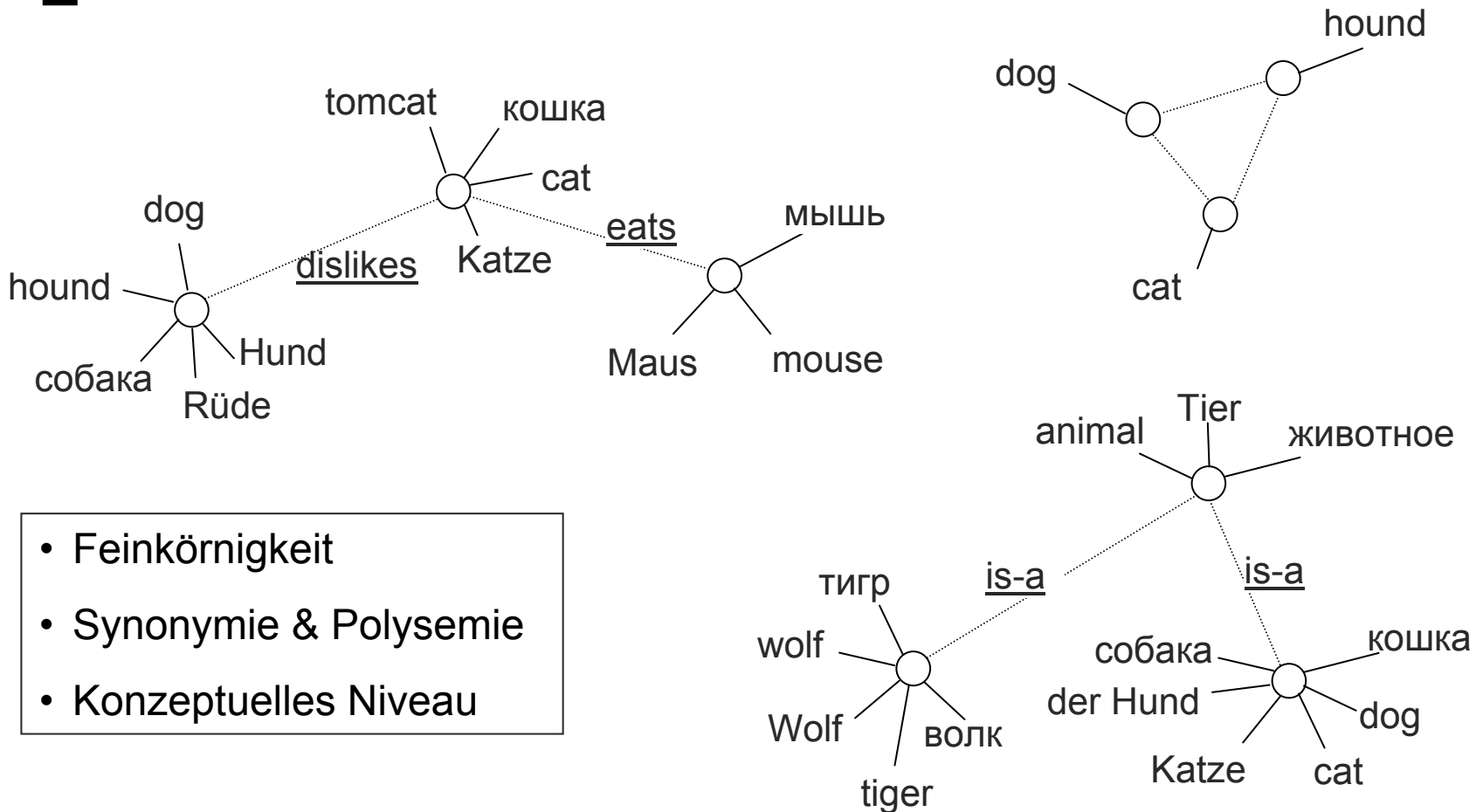
Schichtenarchitektur der Metasuchmaschine

1. Eigenschaften von Suchmaschinen (Protokolle, Anfragesyntax, Antwortmuster usw.)
2. Allgemeine Systemlogik
3. Eigenschaften der Domäne (Thesaurus + Sortierverfahren)

Haupteigenschaften des Thesaurus

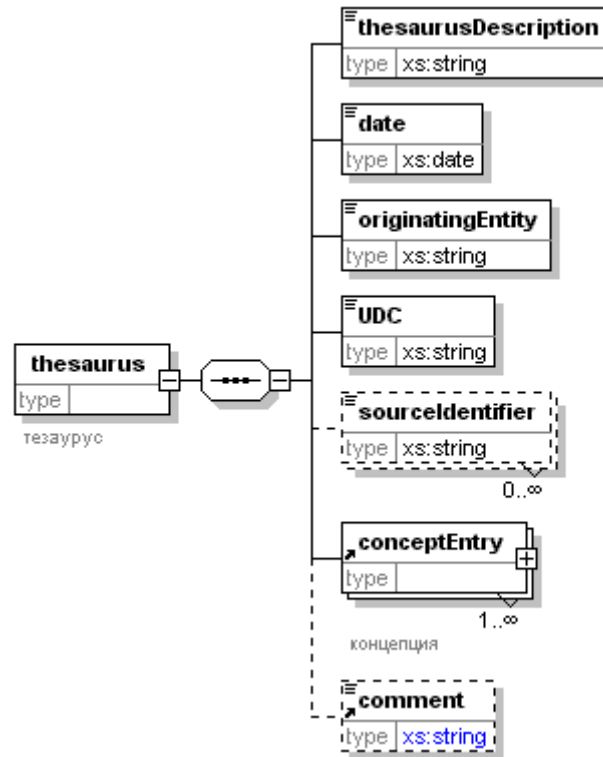
- Hauptelement – die Konzept
(nicht der Terminus)
- Gebietspezifische Relationstypen

Beispiele des Ansatzes

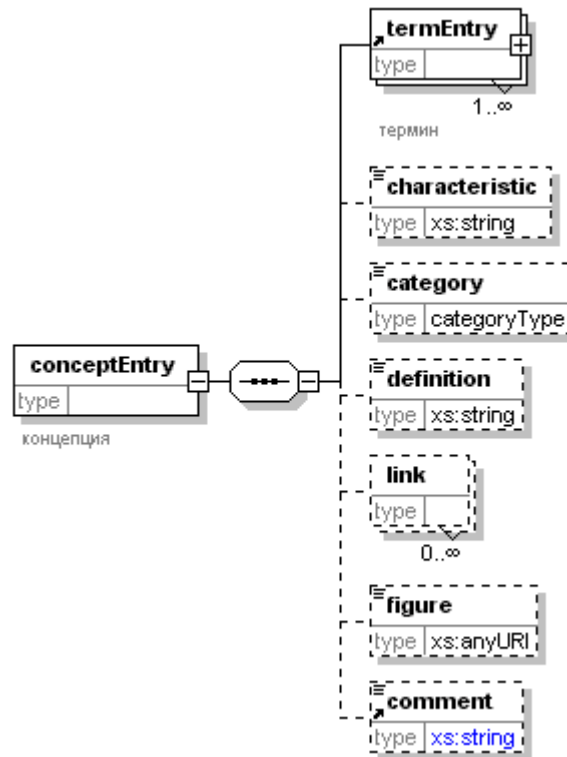


- Feinkörnigkeit
- Synonymie & Polysemie
- Konzeptuelles Niveau

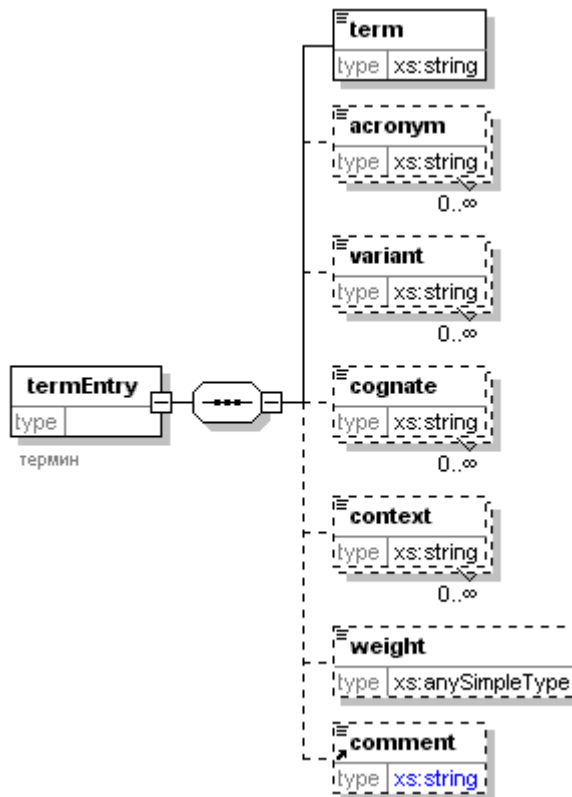
thesaurus Element



conceptEntry Element



[*termEntry* Element]



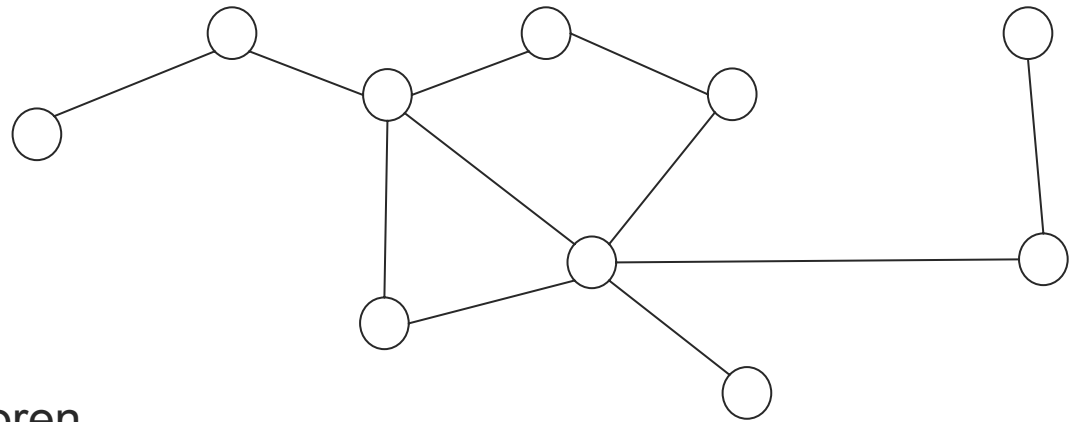
Beispiel für einen Thesaurus

- Anwendungsdomäne:
Automatic Optical Inspection of Printed
Circuit Boards
- ca. 200 Konzepte
- ca. 700 Termini
(inkl. Variante, Abkürzungen usw.)
- ca. 800 bidirektionale Verweise (links)

[Anfrageformulierung]

- Beerenlesemetapher
- Automatische Methoden
 - Musterbasierte Anfragebildung
 - Der kürzeste Weg zwischen zwei Konzepten
 - Anfragelockerung

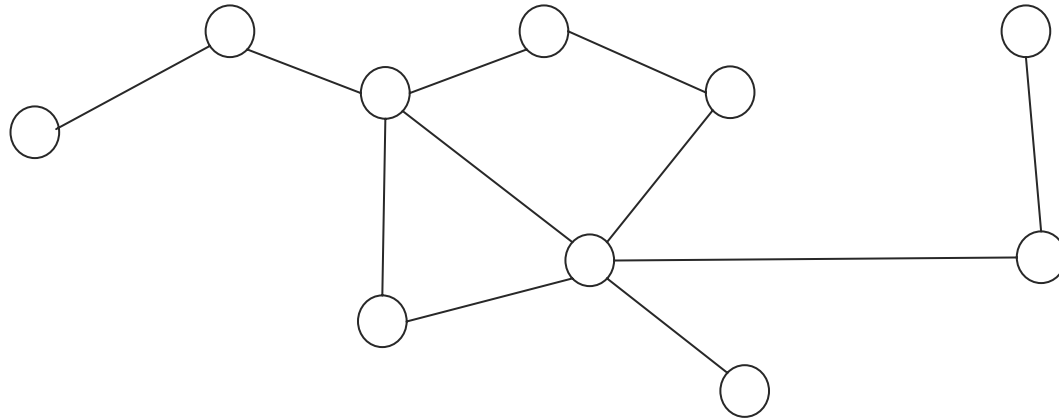
Musterbasierte Anfragebildung



Muster:

1. Relationstypen mit entsprechenden Operatoren (**AND**, OR, ANDNOT)
2. Tiefe der Traversierung
3. Sprachoptionen

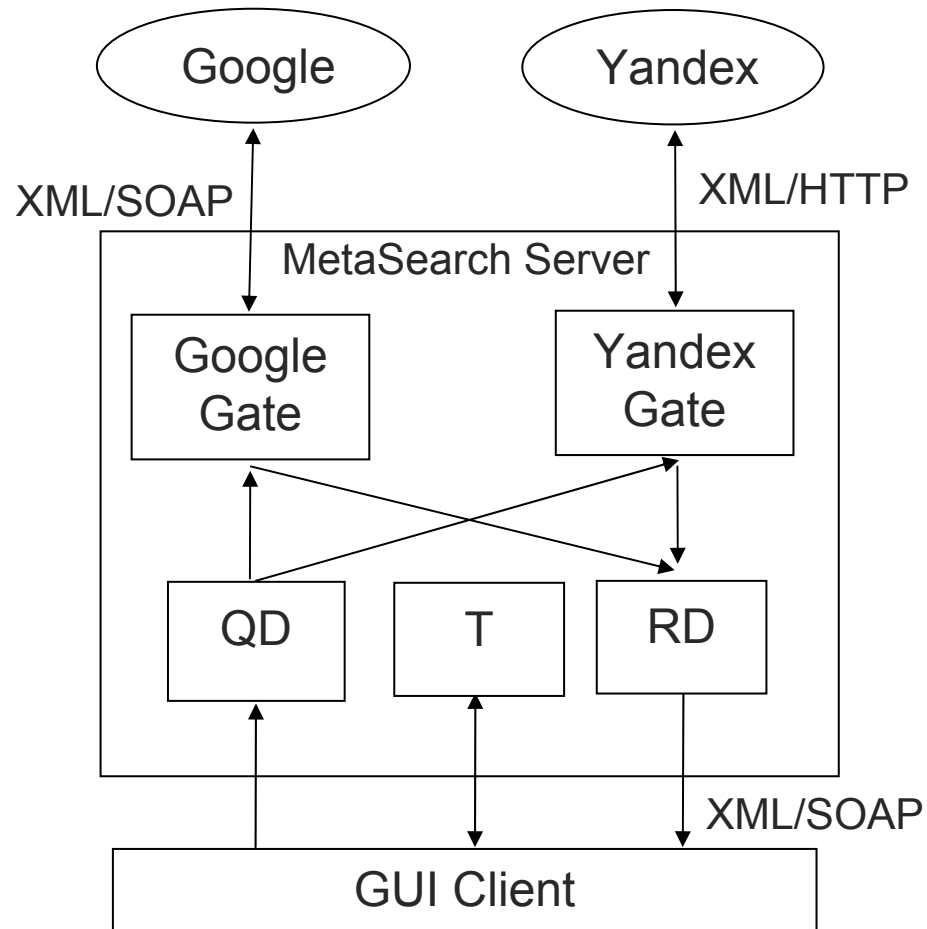
Der kürzeste Weg zwischen zwei Konzepten



Bildung einer Rangliste

- URL (z. B. aus einem Katalogabschnitt)
- Ursprüngliche Position
- Rating der Suchmaschine
- Dateiformat
- Dateigröße
- Datum

ProThes: Allgemeine Systemarchitektur



Entwicklungsrichtung?

- Metasuchmaschine im Internet
- Benutzerschnittstelle zu heterogenen Informationsressourcen
- Semantische Modellierung des Anwendungsbereichs