

# Regionales Rechenzentrum für Niedersachsen, Hannover (RRZN)

## – Informationssysteme –

### Vom Urwald der Wörter zur strukturierten Suche ... Suchhilfe für Informationssysteme und Suchmaschinen

Claus-Peter Rückemann  
05. Juni 2003

25. DGI-Online-Tagung, Competence in Content/comInfo, Frankfurt a. M.

#### **URL (dieser Text/RRZN):**

<http://www.unics.uni-hannover.de/cpr/x/publ/2003/beitrag>  
<http://www.rrzn.uni-hannover.de/>

#### **Kurzfassung:**

Das Regionale Rechenzentrum für Niedersachsen (RRZN) entwickelt und betreibt seit vielen Jahren eine Reihe von verschiedenen Informationssystemen und Suchmaschinen zu unterschiedlichen Zwecken. Die inhaltsbezogene Unterstützung des Nutzers bei der Recherche bzw. Suche mit Informationssystemen, Datenbanken oder Expertensystemen und insbesondere mit Suchmaschinen und externem Datenmaterial im WWW, ist eine komplexe Aufgabe, die im Hintergrund einigen Aufwand erfordert. Neben verschiedenen Möglichkeiten, die an anderen Stellen diskutiert werden ([1], [2], [3], [4], [5]), wird hier kurz und exemplarisch die Möglichkeit für die Nutzung von Thesauri zur Bereitstellung von Suchvorschlägen vorgestellt.

Heim

Titelseite

Letzte Seite



Zurück

Seite 1 von 11

Vollbild

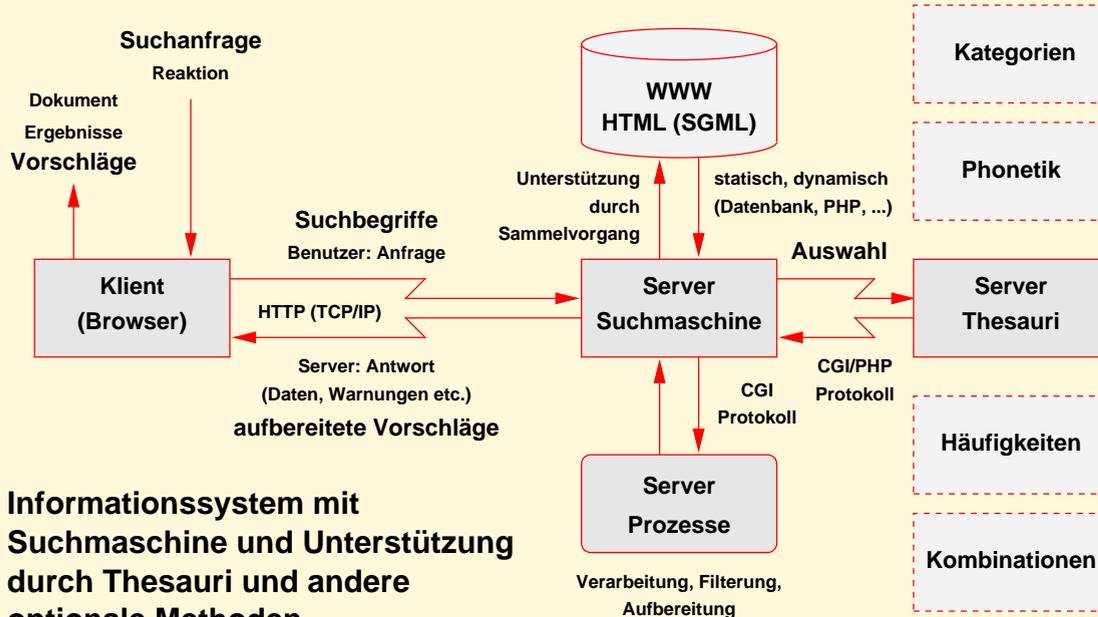
Beenden

Verlassen

## 1. Zusammenspiel Informationssystem – optionale Methoden

### 1.1. Übersicht zum Ablauf

Folgende Übersicht beschreibt eine einfache Nutzung von Thesauri mit einem Informationssystem (Suchmaschine, Recherchesystem, ...) von der Suchanfrage bis zur Rückgabe von Vorschlägen für die Suche.



**Informationssystem mit Suchmaschine und Unterstützung durch Thesauri und andere optionale Methoden**

**Abbildung:** Informationssystem mit Suchmaschine und Unterstützung durch Thesauri und andere optionale Methoden.

Zusammenspiel ...

Anforderungen aus ...

Beispiele

Datenbank und ...

Einschränkungen ...

Heim

Titelseite

Letzte Seite



Zurück

Seite 2 von 11

Vollbild

Beenden

Verlassen

## 1.2. Ablauf

Die wesentlichen Schritte können folgendermaßen zusammengefaßt werden:

1. Suchanfrage, Übermittlung von Daten (u.a. Suchbegriffe) mittels eines Klienten an die Suchmaschine.
2. Verarbeitung, Filterung, Aufbereitung der Anfrage und Abarbeitung damit verbundener Prozesse für die Bearbeitung der Suchanfrage.
3. Ermittlung und Aufbereitung der Ergebnismenge (Daten aus dem WWW, Datenbanken, usw.).
4. Übermittlung bestimmter, ausgewählter Daten (Suchbegriffe) an die Thesauri-Schnittstelle.
5. Verarbeitung der Suchbegriffe auf der Thesauri-Seite.
6. Ermittlung der geeigneten Thesauri-Daten (Sachdaten, Objektdaten, Namen, ...) und Übermittlung an die Suchmaschine.
7. Verarbeitung, Filterung, Aufbereitung der Thesauri-Daten und Integration in die Ergebnismenge.
8. Übermittlung der Ergebnismenge mit den unterstützenden Daten (Thesauri-Daten, ...) an den Nutzer.

Zusammenspiel ...

Anforderungen aus ...

Beispiele

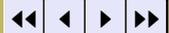
Datenbank und ...

Einschränkungen ...

Heim

Titelseite

Letzte Seite



Zurück

Seite 3 von 11

Vollbild

Beenden

Verlassen

## 2. Anforderungen aus dem Blickwinkel der Suchmaschine

- Gewährleistung einer durchgehenden Verfügbarkeit der eingesetzten Technik. Ausfallsicherheit. Aus diesem Grund werden solche optionalen Methoden nur auf eine Weise in bestehende Suchmaschinen integriert, daß der weitere Betrieb der Suchmaschine, beispielsweise bei einer externen Störung, nicht beeinträchtigt wird.
- Datenschutz.
- Einfache, intuitive Bedienbarkeit.
- Umfang/Länge der Ergebnisse darf Ergonomie und Übersichtlichkeit nicht beeinträchtigen.
- Möglichst hohe Geschwindigkeit bei Ausführung aller Teilabläufe, kurze Latenzzeiten.
- Bestimmte Modalitäten für Abfragevorgänge.
- Bestimmte Beschränkungen (beispielsweise: erlaubte Zeichen, Anzahl und Umfang der Ergebnisse).
- Einsatz von offen und frei verfügbaren und dokumentierten Standards, Methoden, Schnittstellen (CGI, SGML/XML/HTML, PHP, GPL-Werkzeuge, GNU-Linux usw.).
- Orientierung für den Einsatz an langjährigen Erfahrungswerten.

Zusammenspiel...

Anforderungen aus...

Beispiele

Datenbank und...

Einschränkungen...

Heim

Titelseite

Letzte Seite



Zurück

Seite 4 von 11

Vollbild

Beenden

Verlassen

### 3. Beispiele

#### 3.1. Integration in Ausgabeseite

Eine Unterstützung durch Suchbegriffe aus den Thesauri von Axel Springer Verlag Infopool exemplarisch integriert in die **Ausgabeseite der Suchmaschine MetaGer (Live-Link)** des RRZN. Alle folgenden Beispiele entstammen diesen Thesauri. Die Relationen und der Umfang wurden für den speziellen Anwendungszweck reduziert.

**MetaGer**, die Suchmaschine über deutschsprachige Suchmaschinen

Ein Service des **RRZN** & **RVS**: Regionales RechenZentrum für Niedersachsen & Lehrgebiet Rechnernetze und Verteilte Systeme, **Universität Hannover** - **Kontakt**

Suchbegriffe erweitern/verändern oder völlig neue Suche:

**Neue Suche starten**

[Neue/treffendere Suchbegriffe assoziieren?](#)

Folgende Suchvorschläge (HB=Hauptbegriffe) zu Ihrer Abfrage mit freundlicher Unterstützung von AXEL SPRINGER VERLAG INFOPOOL

**Suchbegriff: antarktis** HB: antarktische Oberbegriffe: polargebiete Unterbegriffe: amundsen+scott, britisches+antarktischegebiet, dumont+d+urville, faraday, filchner-station, französisches+antarktischegebiet, georg+forster, georg+von+neumayer, hudson+mountains, macmurdy, mawson-station, mount+berlin, mount+erebus, port+lockroy, südpol, vaughan, vostok, woxcester+range, wostoksee

**Suchbegriff: eis** HB: eis mehrfache - HB: speiseeis, vereisung

**Suchbegriff: seerettung** HB: seerettung Oberbegriffe: rettung, schiffahrt, unglück

**Ausgewählter Vorschlag: ANTARKTIS+EIS+SEERETTUNG**  
Zu Ihrem Suchbegriff konnte automatisch kein alternativer Suchvorschlag ausgewählt werden.  
Bitte wählen Sie ggf. aus obiger Liste die gewünschten Suchbegriffe.

Keine Treffer gefunden :-)

**Abbildung:** Ausgabeseite der Suchmaschine MetaGer des RRZN – ausnahmsweise ohne Treffer (gekürzte Montage), dafür mit Vorschlägen aus den Thesauri von ASV Infopool.

### 3.2. Aufrufe über eine Schnittstelle

Als Vorschläge können sehr flexibel beispielsweise Oberbegriffe zu einem Suchbegriff verwendet werden. Wenn diese nicht vorhanden sind, können Synonyme verwendet werden usw. Auf diese Weise kann die Suche auch inhaltlich breiter und flexibler gestaltet werden.

#### **Abfrage zum Suchbegriff: stern**

**Synonyme:** fixstern

**Oberbegriffe:** himmelskörper

**Unterbegriffe:** 18+scorpii, 1st+tw+hya+association, 51+pegasi, 55+cancri, 8405+asbolus, beta+pictoris, blue+straggler, capella, cw+leonis, delta+velorum, denis-p, eta+carinae, fg+sagittae, fomalhaut, hamburger+stern, he+107-5240, hyaden, kh+15d, mira, nova, p+cygni, pistolenstern, planetarischer+nebel, polarstern, psr+2224+65, pulsar, r+coronae+borealis, r+doradus, red+rectangle-nebel, rho+cas, roter+riese, roter+überriese, roter+zwerg, s50, sakurai, schwarzes+loch, sirius, sonne, supernova, supernova+1987a, supernova+1996x, supernova+1999am, supernova+rxj+0852.0-4622, tau+ceti, tmr-1c, v+338+ser, vla+1623, weißer+zwerg, wolf-rayet+104

**verwandte-Begriffe:** planet

**Eigennamen-HB:** stern

**Eigennamen-UB:** stern+tv+magazin

**Objektthesaurus-HB:** stern

**Objektthesaurus-OB:** fremdobjekt, onlinenutzung+extern, seriöse+presse

**Objektthesaurus-UB:** stern+tv+magazin

[Zusammenspiel...](#)

[Anforderungen aus...](#)

[Beispiele](#)

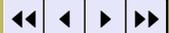
[Datenbank und...](#)

[Einschränkungen...](#)

[Heim](#)

[Titelseite](#)

[Letzte Seite](#)



[Zurück](#)

Seite 6 von 11

[Vollbild](#)

[Beenden](#)

[Verlassen](#)

**Abfrage zum Suchbegriff: planet****Synonyme:** wandelstern**Unterbegriffe:** 1999+cg9, epsilon+eridani, erde+planet, hd+141569, hr+4796a, jupiter, mars, merkur, mond, neptun, nereide, pluto, saturn, uranus, venus**verwandte-Begriffe:** asteroid, komet, stern**Eigennamen-HB:** planet**Eigennamen-mehrfache-HB:** planet+computerpresse, planet+messe, planet+schiff**Abfrage zum Suchbegriff: versicherung****Synonyme:** versicherungsgesellschaft, versicherungsleistung, versicherungsschutz**Unterbegriffe:** bauleistungsversicherung, berufsunfähigkeitsversicherung, edv-versicherung, führerscheinentzugsversicherung, gebrauchtwagenversicherung, graffitiversicherung, gruppenversicherung, jagdversicherung, kinderversicherung, kraftfahrzeugversicherung, kreditversicherung, lottoversicherung, lösegeldversicherung, luft-+raumfahrtversicherung, luxusversicherung, managerversicherung, medienversicherung, patentversicherung, produktschutzversicherung, projektversicherung, rechtsschutzversicherung, reiseversicherung, rentenversicherung, reparaturkostenversicherung, rückrufkostenversicherung, rückversicherung, sachversicherung, seeversicherung, sozialversicherung, speditionsversicherung, tierversicherung, umweltversicherung, unfallversicherung, veranstaltungsversicherung, versicherungsbetrug, versicherungsfremde+leistung, vertrauensschaden-versicherung**verwandte-Begriffe:** schadenersatz, schadensfall*Zusammenspiel...**Anforderungen aus...***Beispiele***Datenbank und...**Einschränkungen...*[Heim](#)[Titelseite](#)[Letzte Seite](#)[Zurück](#)[Seite 7 von 11](#)[Vollbild](#)[Beenden](#)[Verlassen](#)

### Abfrage zum Suchbegriff: essen

**Oberbegriffe:** ernährung

**Unterbegriffe:** essgewohnheit, essstörung, essen

**mehrfache-HB:** bad+essen, essen+belgien, essen+nordrhein-westfalen

### Abfrage zum Suchbegriff: moor

**Oberbegriffe:** landschaft

**Unterbegriffe:** flachmoor, hochmoor, torf, übergangsmoor

**verwandte-Begriffe:** feuchtgebiet, sumpf

### Abfrage zum Suchbegriff: fenn

**HB:** hochmoor

### Abfrage zum Suchbegriff: torf

**Oberbegriffe:** moor

### Abfrage zum Suchbegriff: reise

**Synonyme:** reiseverkehr

**Unterbegriffe:** dienstreise, geschäftsreise, klassenfahrt,

last+minute, reisebetrug, reisejournalismus, reisender, tourismus

**verwandte-Begriffe:** reisewelle, urlaub

### Abfrage zum Suchbegriff: lob

**mehrfache-HB:** anerkennung

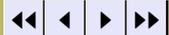
**Synonyme:** kompliment

**Oberbegriffe:** verhalten

Heim

Titelseite

Letzte Seite



Zurück

Seite 8 von 11

Vollbild

Beenden

Verlassen

## 4. Datenbank und Begriffslisten

Um eine Datenbank für den beschriebenen Zweck effizient zu nutzen, muß eine vereinheitlichte Begriffswelt zur Beschreibung der Relationen, Attribute usw. aufgebaut werden. Ein solcher Thesaurus bildet meist eine fachübergreifende Datenbank, die insbesondere die Aufgaben erfüllt, verfügbare Begriffe zur Beschreibung und Dokumentation festzulegen. Sowohl die Strukturierung, als auch die Erstellung von Begriffslisten ist abhängig vom Blickwinkel. In allen Fällen muß je nach Verwendungszweck mit den beteiligten Fachgruppen und Primärnutzern die jeweils beste oder am besten geeignete Lösung gesucht, erstellt und fortlaufend gepflegt werden. Daten zu Bearbeitungen, Datum, Bearbeiter usw. sollten detailliert und kontinuierlich erfaßt werden.

### 4.1. Nutzbarkeit von Begriffslisten

Beispiele für Begriffslisten und deren Nutzbarkeit für die Verwendung mit Suchmaschinen (\* inhaltlicher Pluspunkt, + methodisch-umsetzungstechnischer Pluspunkt):

*	++++	Einfache Begriffsliste
**	+++	Sortierte Begriffsliste
***	++	Hierarchische Begriffsliste
***	++	Netzartige Begriffsliste (Netzwerk von Begriffen)
****	+	Kombinationen einer sortierten und gleichzeitig hierarchischen Begriffsliste
****	+	Kombinationen einer sortierten und gleichzeitig netzartigen Begriffsliste

Daraus läßt sich schlußfolgern, daß es allein schon in dieser Hinsicht – wie bei den meisten Problemen dieser Art – keine optimale Lösung gibt, sondern daß zwischen Aufwand bei der Umsetzung und Aufwand für den Nutzer (z.B. über Menüs geführte Recherche) abzuwägen ist.

Zusammenspiel...  
Anforderungen aus...  
Beispiele  
**Datenbank und...**  
Einschränkungen...

[Heim](#)[Titelseite](#)[Letzte Seite](#)[Zurück](#)[Seite 9 von 11](#)[Vollbild](#)[Beenden](#)[Verlassen](#)

## 5. Einschränkungen für die Nutzbarkeit der optionalen Methoden

- Es ist eine fast vollständige Automatisierung bezüglich des Abfragevorgangs notwendig.
- Das Problem von Mehrfachbedeutungen ist durch vollständige Automatisierung allein nicht zu beheben.
- HTTP ist kein „Gedächtnisprotokoll“ wie z.B. das GLTP (Geographic Locator Transfer Protocol) für räumliche Informationssysteme. Informationen aus möglicherweise zusammenhängenden Nutzungsvorgängen stehen so für sich (Datenschutz).
- Es gibt bei öffentlich zugänglichen Informationssystemen und Suchmaschinen in der Regel keinen suchtechnisch privilegierten Nutzerkreis. Die Abstimmung auf bestimmte Präferenzen der Nutzer ist so sehr schwierig.
- Der Datenschutz für den Nutzer sollte (zumindest bezogen auf den öffentlichen Einsatz und Betrieb) oberste Priorität haben.
- Es ist ein sensibler Umgang bezüglich der Aktivitäten und Kooperationen erforderlich.

### **Konsequenzen für Vorschläge können sich insbesondere ergeben aus:**

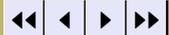
- der Kopplung der Thesauri,
- der Reduzierung der inhaltlichen Konsistenzen durch fehlende Zusatzinformationen,
- dem Aufwand für Pflege und Expertenwissen für Verbreiterung, Vertiefung und Strukturierung der Themengebiete,
- dem jeweiligen Verwendungszweck,
- ... aber beispielsweise auch aus den exemplarisch beschriebenen **Anforderungen** für den Betrieb von Informationssystemen.

Zusammenspiel ...  
Anforderungen aus ...  
Beispiele  
Datenbank und ...  
**Einschränkungen ...**

Heim

Titelseite

Letzte Seite



Zurück

Seite 10 von 11

Vollbild

Beenden

Verlassen

## Literatur

- [1] RÜCKEMANN, C.-P.: *Suchhilfe für Informationssysteme und Suchmaschinen*. [Internet], 2003. URL: <http://www.unics.uni-hannover.de/cpr/x/publ/2003/suchhilfe/suchhilfe.html> (Text).
- [2] RÜCKEMANN, C.-P.: *Informationssysteme und Thesauri*. [Internet], 2003. URL: <http://www.unics.uni-hannover.de/cpr/x/publ/2003/informationssysteme/informationssysteme.html> (Text).
- [3] RÜCKEMANN, C.-P.: *Thesaurus-Unterstützung für Informationssysteme*. [Internet], 2003. URL: <http://www.unics.uni-hannover.de/cpr/x/publ/2003/thesaurus> (Text), URL: <http://metager.de/index0-thesaurus>.
- [4] RÜCKEMANN, C.-P.: *Phonetische Algorithmen zu komparativen Zwecken am Beispiel des phonetischen Komparators für Informationssysteme*. [Internet], 2002. URL: <http://www.unics.uni-hannover.de/cpr/x/publ/2002/phokom> (Text), URL: [http://jserv.rrzn.uni-hannover.de/meta/cgi-bin/cprsndx\\_eqkatll.sh?rueckemann+forschung](http://jserv.rrzn.uni-hannover.de/meta/cgi-bin/cprsndx_eqkatll.sh?rueckemann+forschung).
- [5] RÜCKEMANN, C.-P.: *Kategorien-Lexika zur Nutzung mit Informationssystemen*. [Internet], 2002. URL: <http://www.unics.uni-hannover.de/cpr/x/publ/2002/katlex> (Text).

Zusammenspiel...  
Anforderungen aus...  
Beispiele  
Datenbank und...  
Einschränkungen...

Heim

Titelseite

Letzte Seite



Zurück

Seite 11 von 11

Vollbild

Beenden

Verlassen