

### Terminologiebasiertes Hypertext Retrieval

Willenborg, Josef

Veröffentlichungsversion / Published Version

Sammelwerksbeitrag / collection article

**Empfohlene Zitierung / Suggested Citation:**

Willenborg, J. (1994). Terminologiebasiertes Hypertext Retrieval. In A. Boehm, A. Mengel, & T. Muhr (Hrsg.), *Texte verstehen : Konzepte, Methoden, Werkzeuge* (S. 257-287). Konstanz: UVK Univ.-Verl. Konstanz. <https://nbn-resolving.org/urn:nbn:de:0168-ssoar-14802>

**Nutzungsbedingungen:**

*Dieser Text wird unter einer CC BY-NC-ND Lizenz (Namensnennung-Nicht-kommerziell-Keine Bearbeitung) zur Verfügung gestellt. Nähere Auskünfte zu den CC-Lizenzen finden Sie hier:*

<https://creativecommons.org/licenses/by-nc-nd/4.0/deed.de>

**Terms of use:**

*This document is made available under a CC BY-NC-ND Licence (Attribution-Non Commercial-NoDerivatives). For more information see:*

<https://creativecommons.org/licenses/by-nc-nd/4.0>

# Terminologiebasiertes Hypertext Retrieval

Josef Willenborg  
Technische Universität Berlin

## 1. Einleitung

Es ist bekannt, daß die Suche und Ablage von Information mit Hilfe von Thesauri erleichtert werden kann:

"Es ist unbestritten, daß die Verwendung eines Thesaurus bei der Indexierung und bei der Formulierung von (...) Anfragen die Retrievaleffektivität erheblich steigert."

(Salton, McGill 1987 240)

Wenig erforscht ist hingegen die Idee, das thesaurusbasierte Information Retrieval zum terminologiebasierten Information Retrieval zu erweitern und dessen Qualität und Effizienz durch den Einsatz von Hypertextkonzepten noch weiter zu verbessern. Es ergeben sich die folgenden Vorteile einer solchen Sicht:

- einheitliche Erfassung, Zusammenführung und Austauschbarkeit vorhandener Terminologien (Wiederverwendbarkeit)

"Da die Erstellung von Wissensbasen teuer ist, sollte größeres Augenmerk auf die Wiederverwendbarkeit von Wissensbasen gelegt werden. (...) Die Entwicklung großer Expertensysteme wird mit durchschnittlich einer Million Dollar veranschlagt. (...) Für zahlreiche Anwendungen wären Wissensbasen sehr nützlich, wenn sie billig und in einfacher Form zur Verfügung stünden."

(Czedik 1992)

- Verbesserung der Konsistenz terminologischer Bestände
- Verbesserung der Retrievalergebnisse durch Anreicherung von Terminologien (Multimedialisierung, Multilingualisierung und Hypertextualisierung)
- Verbesserung der Benutzerschnittstelle durch die Kombination einheitlicher, kommando-orientierter Anfragemöglichkeiten (universelle Anfragesprache) mit Navigations- und Browsingaspekten

- terminologiegestützte Information Retrieval Systeme können mit Hilfe neuer objektorientierter Techniken mit verhältnismäßig geringem Aufwand realisiert werden (Software Engineering und Design, Kopplung von Programmiersprache und Datenbank).

In dieser Arbeit wird eine einfache aber mächtige Anfragesprache für hypertextuelle Terminologie- und Informationsbestände aufgebaut. Es werden Sprach-elemente für elementare, zusammengesetzte und gruppierte Informationseinheiten und für n-stellige Relationen zwischen diesen bereitgehalten.

Mit dem Computerprogramm PflaSaurus (Willenborg 1991, Willenborg 1994) wurden bereits Verfahren zur Navigation in Thesauri objektorientiert realisiert. Die Bearbeitung der Bestände wird kooperativ auf unterschiedlichen Formaten durchgeführt. Insbesondere werden Verfahren zur Sicherung der Konsistenz (Mengel 1991) und zur übersichtlichen Darstellung und dem Browsing bereitgestellt.

Insgesamt existiert heute eine große Menge terminologischer Daten für unterschiedliche Fachgebiete und Anwendungen, die sich beim Information Retrieval nutzen lassen. Thesaurus Guide 1985 zählt beispielsweise allein in Europa knapp tausend Thesauri. Noch umfassender ist die Sammlung von Infoterm (Felber, Budin 1989 257).

Es lassen sich die folgenden Arten von Terminologien unterscheiden (Felber, Budin 1989, Schönfeldt 1993, Willenborg 1993). Sie sollen hier unter dem Oberbegriff Terminologie subsumiert und weiterverwendet werden.

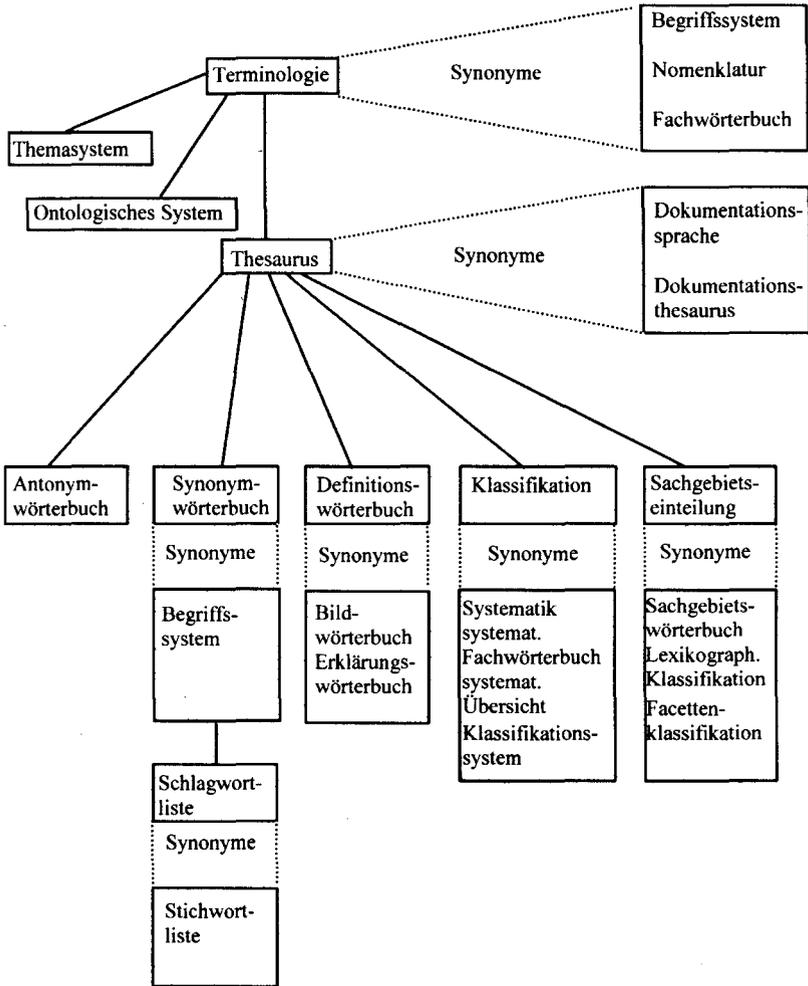


Abb. 1: Terminologiearten.

## 2. Terminologiebasiertes Information Retrieval

Es ist lohnenswert, den Begriff des Information Retrieval unabhängig vom jeweils verwendeten Informationssystem einzuführen.

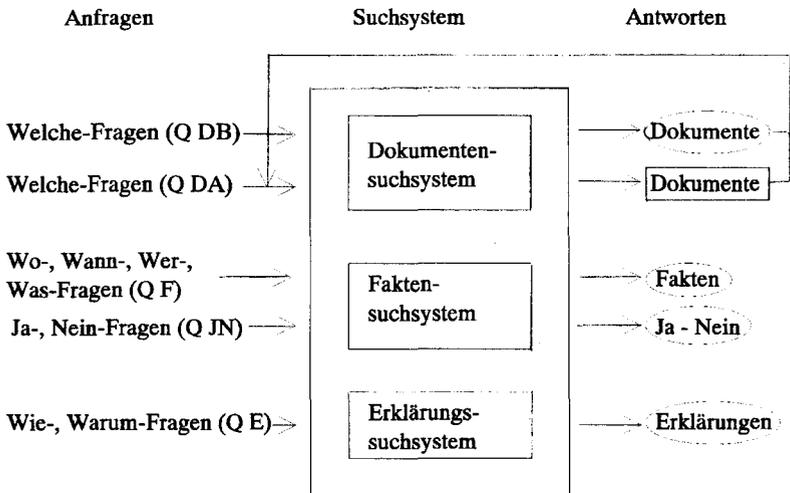
"Von besonderem Interesse sind in diesem Zusammenhang allgemeine Systemschnittstellen, die es dem Nutzer erlauben, seine Suchanfrage unabhängig vom System in einer einheitlichen Suchsprache zu formulieren. Die Systemschnittstelle sollte dann diese allgemeine Suchsprache in die unterschiedlichen Systemsprachen der einzelnen Informationssysteme übersetzen (...). Systemübergreifende Benutzerschnittstellen sind deshalb äußerst nützlich, weil sie dem Nutzer die Mühe abnehmen, sich mit internen Details der einzelnen Informationssysteme auseinanderzusetzen. Von weit größerem Interesse sind Systeme, die selbst entscheiden, welche Informationsquelle bei einer bestimmten Anfrage abzufragen ist. (...) Der Entwurf eines solchen intelligenten Informationssystems ist zur Zeit noch nicht möglich."

(Salton, McGill 1987 454)

Ulrike Reiner (Reiner 1991 35) legt für diesen Bereich den Grundstock und baut mit IQL (Intermediary Query Language) eine universelle Anfragesprache für Informationssysteme auf.

"(...) führen wir den Begriff des 'Informationssuchsystems' (oder kurz: 'Such-system') als Oberbegriff für Dokumenten-, Fakten- und Erklärungssysteme ein. (...)

(...) Integriertes Modell eines Dokumenten-, Fakten- und Erklärungssuchsystems"  
 (Reiner 1991 35)



**Abb. 2:** Integriertes Modell eines Dokumenten-, Fakten- und Erklärungssystems (Reiner 1991 35).

Der Hauptvorteil des terminologiebasierten Information Retrieval besteht darin, daß der Informationssuchende während seiner Suche erkennt, wie die Informationsbestände durch die Terminologien strukturiert sind. Er bekommt terminologische Hilfestellungen angeboten, anhand der er entscheidet, welches terminologische Vokabular in seiner Suche Erfolg verspricht. Der Informationssuchende navigiert beispielsweise graphisch in dem Terminologienetz und wählt Einträge,

die er in seiner Anfrage berücksichtigen will. Der Vorteil dieser terminologiegeführten Suche besteht darin, daß der Anfragende sieht, welche Termini in einem Sachgebiet zur Suche vorhanden sind und welche inhaltlich verwandten Termini existieren.

Definition [1]: Terminologiebasiertes Information Retrieval heißt der Prozeß, bei dem eine Anfrage mit terminologischen Sprachelementen eine Teilmenge des zugrundeliegenden Terminologie- und Informationsbestandes als Ergebnis zurückliefert.

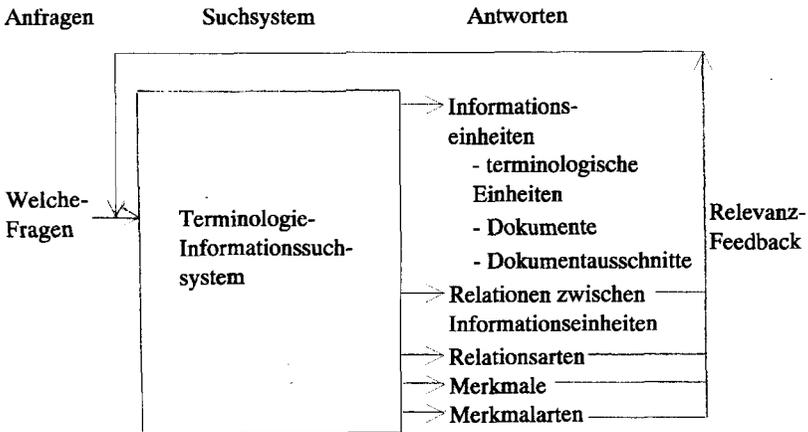


Abb. 3: Terminologie-Informationssuchsystem.

Zur Verbesserung des terminologiebasierten Information Retrieval können Morphologiekomponenten eingesetzt werden. Mittels solch einer Komponente (beispielsweise MOLEX vom IDS Mannheim, das in Saarbrücken entwickelte System MORPHIX, das von Gazdar entwickelte System DATR oder das two-level System KIMMO) wird in einem ersten Schritt die Reduzierung einer Wortform auf ihre Grundform vorgenommen. In einem zweiten Schritt kann dann mit der Terminologiekomponente die Abbildung dieser Grundform auf terminologische Einheiten durchgeführt und so schließlich die Menge der indextierten Informationseinheiten bestimmt werden:

Anfragetermini

|  
| Morphologiekomponente

Grundform(en)

|  
| Terminologiekomponente

terminolog. Einträge ————— (terminolog.) Informationseinheiten

**Abb. 4:** Information Retrieval mit Morphologie- und Terminologiekomponente.

Das Wechselspiel kommandoorientierter Anfragenformulierung und der Navigation in Informationsbeständen stellt einen wichtigen Forschungsgegenstand dar. Beispielsweise kann die Strategie gewählt werden, mit einem Anfragekommando einen Einstiegspunkt in die informationelle Verweisstruktur zu gewinnen, um dann mit Navigationsmethoden die Informationmenge weiter einzuschränken. Die umgekehrte Strategie besteht darin, mit Hilfe von Navigationswerkzeugen ein Gefühl für die Begrifflichkeiten des Fachgebietes zu entwickeln, um dann mit gezielten Anfragekommandos die gewünschte Information abzurufen.

Die Verfahren zur Selektion, Auswertung, Analyse und Weitergabe gefundener terminologischer Information sind dabei einzubeziehen. Dem Informationssuchenden sollte bei seiner Anfragenvor- und -nachbereitung Unterstützung angeboten werden. Anfrageergebnisse werden für die zukünftige Suche in den Informationsbeständen festgehalten. Relevanzfeedbackverfahren sind dabei auf deren Einsatz zu prüfen (Salton, McGill 1987 255).

Weiterhin werden häufig verwendete Suchtermini erfaßt und dem Informationssuchenden als Kandidaten zur Verfügung gestellt. Zusätzlich können die Verwendungshäufigkeiten der Anfragetermini in dem Information Retrieval System als externe Kandidaten berücksichtigt werden.

Für die Benutzeroberfläche eines Information Retrieval Systems ist die Form der Ergebnispräsentation besonders wichtig. Je ausführlicher die Suchergebnisse dargestellt werden, desto leichter ist es für den Nutzer, die nachgewiesene Information hinsichtlich ihrer Relevanz zu beurteilen, desto mehr Zeit kostet jedoch deren Sichtung.

### 3. Terminologien als Hypertext

Terminologien als Hypertext, also als Netzwerke aus (typisierten) Knoten (Texte, Graphik oder multimediales Material) und (typisierten) Kanten zu betrachten, ist erst in letzter Zeit in den Vordergrund der Terminologieforschung gerückt (McMath, Rada, Tamaru 1989, Pollard 1990, Fischer, Möhr 1991, Fowler, Fowler, Wilson 1991, ISO 1087, ISO 10744, ISO 12200, ISO 12620).

Die folgenden Definitionen zur Beschreibung hypertextueller Terminologie- und Informationsbestände werden aus Willenborg 1993 übertragen.

#### **Definition [1]:**

Ein hypertextueller Terminologie- oder Informationsbestand ist eine Menge von:

1. Informationseinheiten
2. Typen von Informationseinheiten
3. Relationen
4. Relationsarten
5. mathematischen Typen von Relationsarten
6. semantischen Typen von Relationsarten
7. Merkmalarten
8. Merkmalen

#### **Definition [2]: Informationseinheit**

Eine Informationseinheit wird wie bei (Kuhlen 1991 333) definiert. Sie besitzt ein Feld zur eindeutigen Kennung (Name, Titel, Ikone), ein Feld für den Typ dieser Einheit, ein Feld für den Inhalt der Einheit, ein Feld für die Merkmalarten und Merkmale, ein Feld für das Sprachenzeichen (Landessprache), ein Feld für den zuständigen Pfleger, ein Feld für den Zeitpunkt der Erzeugung und ein Feld für die externe Quelle.

Eine Informationseinheit kann elementar, zusammengesetzt oder gruppiert aufgebaut werden. Zusammengesetzte Informationseinheiten bestehen aus den elementaren oder wieder aus zusammengesetzten Informationseinheiten. Gruppen fassen Informationseinheiten zu Mengen zusammen.

### **Definition [3]: Typ einer Informationseinheit**

Ein Typ einer Informationseinheit ist entweder elementar, zusammengesetzt, gruppiert oder terminologiespezifisch. Er besitzt ein Feld für die Layout-Eigenschaft, den zuständigen Pfleger und ein Feld für den Zeitpunkt der Erzeugung.

### **Definition [3.1]: Elementarer Typ**

Ein elementarer Typ ist: INTEGER, CHAR, STRING, TEXT, GRAPHIK, BILD, AUDIO, GRAPHIK-ANIMATION, VIDEO,...

### **Definition [3.2]: Zusammengesetzter Typ**

Ein zusammengesetzter Typ ist eine geordnete Menge von elementaren, oder wieder zusammengesetzten Typen. Er besitzt ein Feld zur eindeutigen Kennung (Name, Titel, Ikone).

### **Definition [3.3]: Gruppentyp**

Ein Gruppentyp ist eine Menge von elementaren, zusammengesetzten oder wieder gruppierten Typen. Er besitzt ein Feld zur eindeutigen Kennung (Name, Titel, Ikone).

### **Definition [3.4]: Terminologiespezifischer Typ**

Ein terminologiespezifischer Typ ist: GEGENSTAND, BEZEICHNUNG, BEGRIFF, THEMA, KONTEXT, DEFINITION, ERKLÄRUNG, ANMERKUNG, BEISPIEL, FORMEL, PHONETISCHE UMSCHRIFT, HOMONYM, ...

### **Definition [4]: Relation**

Eine Relation wird zwischen Informationseinheiten beliebigen Typs aufgebaut. Sie besitzt ein Feld für die Relationsart, ein Feld für die Argumentwerte (Informationseinheiten), ein Feld für den zuständigen Pfleger und ein Feld für den Zeitpunkt der Erzeugung.

**Definition [5]: Relationsart**

Eine Relationsart spezifiziert eine Relation. Sie besitzt ein Feld zur eindeutigen Kennung (Name, Titel, Ikone), ein Feld für die erlaubten Argumente (Typen von Informationseinheiten), ein Feld für die Menge der mathematischen Typen der Relationsart, ein Feld für die Menge der semantischen Typen der Relationsart, ein Feld für das Sprachenzeichen (Landessprache), ein Feld für die Menge der Layout-Eigenschaften der Relationsart, ein Feld für den zuständigen Pfleger und ein Feld für den Zeitpunkt der Erzeugung.

Relationsarten sind (hier die eindeutige Kennung durch den Namen):

Ontologische Relationen (Gegenstandsrelationen):

Bestandsrelation (räumlich), Nacheinander-Relation (zeitlich), Relation zum Stoffgebilde, Wirk-Relation, ...

Begriffliche Relationen:

Ober/Unterbegriffsrelation, Verwandtschaftsrelation, Antonymie, Logische Nebenordnung, Logische Diagonalbeziehung, Logische Begriffsleiterbeziehung, Logische Reihenbeziehung,...

Relationen zur Begriffsbeschreibung und Erläuterung:

Inhaltsbeschreibung (Definition), Umfangsbeschreibung, Beispiel, Abbildung, Erläuterung, Kommentar, ...

Themarelationen:

Ober/Unterthema, verwandtes Thema, Themaleiter, Themareihe, Überschneidung, ...

Relationen zwischen Begriffszeichen (Benennung) und Begriff:

Homonymie, Synonymie, Quasi-Synonymie, Wortform, Langform, Kurzform, ...

Übersetzungsrelationen:

Deutsch-Englisch, Deutsch-Japanisch, ...

Andere Relationen:

Produzent-Produkt, Sender-Empfänger, ...

**Definition [6]: Mathematischer Typ einer Relationsart**

Ein mathematischer Typ einer Relationsart ist: REFLEXIV, IRREFLEXIV, SYMMETRISCH, ASYMMETRISCH, ANTISYMMETRISCH, TRANSITIV, ATRANSITIV, ÄQUIVALENT, LINKSEINDEUTIG, RECHTSEINDEUTIG, EINEINDEUTIG, LINKSTOTAL, RECHTSTOTAL, FUNKTION, INJEKTIV, SURJEKTIV, BIJEKTIV, ZYKLISCH, AZYKLISCH, ...

**Definition [7]: Semantischer Typ einer Relationsart**

Ein semantischer Typ einer Relationsart ist: ASSOZIATIV, SYNTAGMATISCH, ANNOTATIV, QUERY LINK, COMMUNICATION LINK, EXECUTABLE LINK, REFERENTIELLER LINK, EXTRA REFERENTIELLER LINK, TYPISIERT, SEMANTISCH, HIERARCHISCH, POLYHIERARCHISCH, POLYDIMENSIONAL, KONZEPTUELL, KOHÄRENZ BEZOGEN, PRÄDIKATIV ARGUMENTATIV, THEMATISCH, RHETORISCH, ILLOKUTIONÄR, ...

**Definition [8]: Merkmalart**

Eine Merkmalart ist ein Gesichtspunkt, nach dem eine Informationseinheit unterteilt wird.

Eine Merkmalart ist: FORM, GRÖSSE, STOFF, HÄRTE, FARBE, VERWENDUNG, RÄUMLICHE EINGLIEDERUNG, HERSTELLER, ERFINDER, HÄNDLER, HERSTELLVERFAHREN, HERKUNFTSLAND, HERKUNFTSZEIT, ...

**Definition [9]: Merkmal**

Ein Merkmal ist eine eindeutige Eigenschaft, die einer Informationseinheit anhand einer Merkmalart zugesprochen wird.

Die semantischen Typen der Relationsarten sind an Kühlen 1991 angelehnt worden, jedoch um die Typen POLYHIERARCHISCH und POLYDIMENSIONAL erweitert worden.

Die Relationsarten in Definition [5] werden überwiegend aus Felber, Budin 1989 übertragen. Beispiele für die Definitionen [1] bis [9] finden sich in Willenborg 1993.

## 4. Mathematische Eigenschaften terminologischer Relationsarten

Mathematische Eigenschaften von Relationsarten lassen sich zur Konsistenzsicherung von Terminologie- und Hypertextbeständen nutzen. Die Konsistenz der Bestände wird anhand der mathematischen Eigenschaften und der Argumenttypen der Relationsarten überprüft. Sie wird dadurch unabhängig von dem Namen und der intendierten aber nicht festgelegten Semantik der Relationsarten.

Es gilt weiterhin allgemein:

1. Jede terminologische Relationsart besitzt eine inverse Relationsart (Vermeidung von Fallgruben).
2. Jede symmetrische terminologische Relationsart besitzt als inverse Relationsart sich selber.
3. Es existieren bezüglich einer Relationsart keine Mehrfachkanten zwischen zwei Knoten.
4. Für jede hierarchische, polyhierarchische oder polydimensionale Relationsart gilt die Azyklizität und die Atransitivität.

In (Schönfeldt 1993) wird die Anwendung mathematischer Eigenschaften auf terminologische Relationsarten untersucht. Dabei lassen sich zusammengefaßt die folgenden Eigenschaften unterscheiden:

Relationsart	refl.	irrefl.	symm.	anti-symm.	trans.	a-trans.
Synonymie	ja	nein	ja	nein	ja	nein
Quasi-Synonymie	ja	nein	ja	nein	nein	nein
Abstraktionsrelation	nein	ja	nein	ja	ja	nein
Bestandsrelation	nein	---	nein	ja	---	nein
Functional component	nein	nein	nein	ja	nein	nein
Segmented whole	nein	ja	nein	ja	ja	nein
Collection member	nein	ja	nein	ja	nein	nein
Set-subset	nein	nein	nein	ja	ja	nein

Logische Gleichordnung (monohierarchisch)	nein	nein	ja	nein	ja	nein
Logische Gleichordnung (polyhierarchisch)	nein	nein	ja	nein	nein	nein
Räumliche Nebenordnung	nein	nein	ja	nein	nein	nein
Kausalbeziehung (direkt)	nein	ja	nein	ja	nein	ja
Kausalbeziehung (indirekt)	nein	ja	nein	ja	ja	nein
Nachfolgebeziehung (direkt)	nein	nein	nein	nein	nein	ja
Nachfolgebeziehung (indirekt)	nein	nein	nein	nein	ja	nein
Antonymie	nein	ja	ja	nein	nein	ja

Tab. 1: Mathematische Eigenschaften terminologischer Relationsarten

Außer den hier betrachteten Eigenschaften lassen sich den terminologischen Relationsarten weitere in Definition [6] genannte mathematische Eigenschaften zuordnen. Diese Untersuchung soll jedoch nicht Gegenstand dieser Arbeit sein.

## 5. HyQL: Eine prädikatenlogische Anfragesprache für hypertextuelle Terminologie- und Informationsbestände

Der Zugriff auf Information mit Hilfe von Terminologien bereitet vor allem aus dem Grunde Schwierigkeiten, daß die in den Fachgebieten erstellten Terminologie- und Informationsbestände in unterschiedlichen Formaten vorliegen, deren Interpretationen sich stark voneinander unterscheiden. Qualitativ hochwertiges und effektives terminologiebasiertes Information Retrieval verlangt deshalb die Entwicklung einheitlicher und übergreifender Anfrageverfahren, die die Vielfältigkeit und Komplexität der Bestände berücksichtigen.

Erster Schritt bei der Entwicklung einer solchen Anfragesprache ist die formale Beschreibung der zugrundeliegenden Terminologie- und Informationsbestände (Willenborg 1993). In einem zweiten Schritt wird die Anfragesprache mit Hilfe einer Methode aufgebaut, die eine scharfe Trennung von Syntax und Semantik

vorsieht (Reiner 1991): Welches Vokabular steht für eine Anfrage zur Verfügung und in welcher Kombination kann es benutzt werden (Syntax)? Was bedeutet eine syntaktisch korrekt gebildete Anfrage bzw. was liefert sie als Ergebnis zurück? Wie ist also das Verhältnis der Anfragezeichen zur Wirklichkeit definiert (Semantik)?

"Wenn nämlich erreicht werden kann, eine Sprache syntaktisch und semantisch getrennt aufzubauen, gilt unter bestimmten Voraussetzungen, daß sich alle Folgerungen aus einem Axiomensystem algorithmisch (syntaktisch) gewinnen lassen."

(Reiner 1991 10).

Je umfangreicher und je komplexer die Syntax und Semantik einer Sprache aufgebaut ist, desto schwieriger wird deren Erlern- und Beherrschbarkeit. Die Anfragesprache HyQL (Hypertextual Query Language) ist als Sprache zwischen Mensch und Maschine konzipiert und nur für den Experten direkt benutzbar. Sie stellt die Grundlage für die Navigations- und Darstellungsmittel eines Informationssystems dar, für geübte und auch ungeübte Benutzer (Reiner 1991 180).

Ähnliche Ansätze zur Beschreibung von Anfragesprachen für Hypertexte finden sich in (Afrati, Koutras 1990) und (Beeri, Kornatzky 1990).

HyQL ist eine prädikatenlogische Sprache erster Stufe. Die Semantik wird modelltheoretisch aufgebaut.

HyQL enthält Sprachelemente für elementare, zusammengesetzte und Gruppen von Informationseinheiten und für Relationsarten. Es können n-stellige Relationen zwischen den Informationseinheiten selber, aber auch zwischen Teilen von Informationseinheiten (bei Conklin 1987 34 Regionen genannt, Informationseinheiten können dabei überlappen) aufgebaut werden.

Weiterhin werden Operatoren eingeführt, die sich auf die von einer Informationseinheit erreichbaren Informationseinheiten beziehen: Vorfahren, Nachfahren und Geschwister (Reiner 1991 121).

Relationsarten werden in HyQL als Individuen anstatt als Prädikate eingeführt, damit Erweiterungen der Sprache (zum Beispiel um Eigenschaften von Relationsarten) vorgenommen werden können, ohne eine höhere Stufe der Prädikatenlogik zu erreichen.

Anfragen werden mit dem Lambda-Operator  $\lambda$  und den Formeln gebildet. Damit können Extensionen von Prädikaten betrachtet werden.

## Syntax

### Alphabet

1. Elementare Informationseinheiten (Individuenkonstanten):  $e_{11}, e_{12}, \dots, e_{21}, \dots, e_{nm}$   
 wobei:  $m$  = Anzahl der Individuen für eine elementare Sorte  
 $n$  = Anzahl der elementaren Sorten
2. Informationseinheiten (Individuenvariablen):  $x, x_{11}, x_{12}, \dots, x_{21}, \dots, x_{pm}$   
 wobei:  $p$  = Anzahl der Sorten
3. Relationsarten (Individuenkonstanten):  $r_1, r_2, \dots, r_k$
4. Relationsarten (Individuenvariablen):  $y, y_1, y_2, \dots, y_k$
5. Zusammengesetzte (Kompositionen) und Gruppen von Informationseinheiten (Funktionskonstanten):  $k_1, \dots, k_{q1}$  und  $g_1, \dots, g_{q2}$
6. Operatorsymbole für Nachfahren bis zur Stufe  $\alpha$ , Vorfahren bis zur Stufe  $\beta$ , Nachfahren bis zu den Blättern (alle Nachfahren), Vorfahren bis zu den Wurzeln (alle Vorfahren), alle Nachfahren und Vorfahren, alle Geschwister (Funktionskonstanten):  $\alpha \uparrow, \beta \downarrow, \uparrow \uparrow, \downarrow \downarrow, \uparrow \downarrow, \downarrow \uparrow, \Leftrightarrow$
7. Beziehungen (Prädikatenkonstanten): REL, OP
8. Logische Symbole:  $\neg$  (nicht),  $\wedge$  (und),  $\vee$  (oder),  $\exists$  (es gibt ein),  $\forall$  (für alle),  $\lambda$  (die Menge)
9. Technische Symbole: ( , )

### Terme

1. Elementare Informationseinheiten (Konstanten und Variablen) sind elementare Terme.
2. Wenn  $t_1, t_2, \dots, t_n$  elementare oder zusammengesetzte Terme sind und  $k_i$  ein Funktionssymbol ist, dann ist  $k_i(t_1, t_2, \dots, t_n)$  ein zusammengesetzter Term (Komposition).
3. Wenn  $t_1, t_2, \dots, t_n$  elementare oder zusammengesetzte Terme oder Gruppenterme sind und  $g_i$  ein Funktionssymbol ist, dann ist  $g_i(t_1, t_2, \dots, t_n)$  ein Gruppenterm (Gruppe).
4. Wenn  $t$  ein elementarer oder zusammengesetzter Term oder Gruppenterm ist und  $r, y$  Relationsarten sind, dann sind  $\alpha \uparrow_r^t, \alpha \uparrow_y^t, \beta \downarrow_r^t, \beta \downarrow_y^t, \uparrow \uparrow_r^t, \uparrow \uparrow_y^t, \downarrow \downarrow_r^t, \downarrow \downarrow_y^t, \uparrow \downarrow_r^t, \uparrow \downarrow_y^t, \downarrow \uparrow_r^t, \downarrow \uparrow_y^t, \Leftrightarrow_r^t, \Leftrightarrow_y^t$  Operatorterme.
5. Das sind alle Terme.

**Formeln**

1.  $\text{REL}(t_1, t_2, \dots, t_n, r), \text{REL}(t_1, t_2, \dots, t_n, y)$ ,  
wobei:  $t_i$  sind Terme (außer Operatorterme);  
 $n$  ist die Stelligkeit der Relationsart  $r$  bzw.  $y$   
 $\text{OP}(x, t)$   
wobei:  $t$  ist Operatorterm  
sind Formeln.
2. Wenn  $F, F_1$  und  $F_2$  Formeln sind, dann sind auch  
 $\neg F, (F_1 \wedge F_2), (F_1 \vee F_2), (\exists x_{ij})F, (\forall x_{ij})F, (\exists y_i)F, (\forall y_i)F$  Formeln.
3. Das sind alle Formeln.

**Anfragen**

1. Wenn  $F$  eine Formel ist, dann ist  $(\lambda \text{REL})(\lambda x_{ij})F$  eine Anfrage.  
Wenn  $F$  eine Formel ist, dann ist  $(\lambda \text{REL})(\lambda y_i)F$  eine Anfrage.
2. Das sind alle Anfragen.

**Semantik**

$p$  sei die Anzahl aller Bereiche.

$E_1$  sei eine Menge von Informationseinheiten aus dem Bereich 1.

$E_2$  sei eine Menge von Informationseinheiten aus dem Bereich 2.

$E_p$  sei eine Menge von Informationseinheiten aus dem Bereich  $p$ .

$\Delta = \{E_1, E_2, \dots, E_p\}$ .

$R$  sei eine Menge von Relationsarten.

**Interpretation der nichtlogischen Symbole**

$\mathfrak{I}(e_{1i}) \in E_1; \mathfrak{I}(e_{2i}) \in E_2; \dots; \mathfrak{I}(e_{ni}) \in E_n; \quad (i=1, \dots, m)$

$\mathfrak{I}(x_{1i}) \in E_1; \mathfrak{I}(x_{2i}) \in E_2; \dots; \mathfrak{I}(x_{pi}) \in E_p; \mathfrak{I}(x) \in \Delta; \quad (i=1, \dots, m)$

$\mathfrak{I}(r_i) \in R$ ;  $\mathfrak{I}(y_i) \in R$ ;  $i \in \mathbb{N}$   
 $\mathfrak{I}(k_i) \in \{\Delta \rightarrow \Delta\}$ ;  $\mathfrak{I}(g_i) \in \{\Delta \rightarrow \Delta\}$ ;  
 $\mathfrak{I}(\text{REL}) \subseteq \Delta \times R$

**Terme**

$\mathfrak{I}(x_{ij})$  und  $\mathfrak{I}(e_{ij})$  sind schon erklärt.

$\mathfrak{I}(k_i(t_1, t_2, \dots, t_n)) = \mathfrak{I}(k_i) (\mathfrak{I}(t_1), \mathfrak{I}(t_2), \dots, \mathfrak{I}(t_n))$ .

$\mathfrak{I}(g_i(t_1, t_2, \dots, t_n)) = \mathfrak{I}(g_i) (\mathfrak{I}(t_1), \mathfrak{I}(t_2), \dots, \mathfrak{I}(t_n))$ .

Die Operatorterme werden simultan mit den atomaren Formeln erklärt.

**Formeln**

1.  $\mathfrak{I}$  ist Modell von  $\text{REL}(t_1, t_2, \dots, t_n, r)$  gdw  $\langle \mathfrak{I}(t_1), \mathfrak{I}(t_2), \dots, \mathfrak{I}(t_n), \mathfrak{I}(r) \rangle \in \mathfrak{I}(\text{REL})$ .

$\mathfrak{I}$  ist Modell von  $\text{REL}(t_1, t_2, \dots, t_n, y)$  gdw  $\langle \mathfrak{I}(t_1), \mathfrak{I}(t_2), \dots, \mathfrak{I}(t_n), \mathfrak{I}(y) \rangle \in \mathfrak{I}(\text{REL})$ .

2.  $\mathfrak{I}$  ist Modell von  $\text{OP}(x, \alpha \uparrow^1_r)$  gdw  $\mathfrak{I}$  ist Modell von  $\text{REL}(t, x^1, r)$   
 und  
 $\mathfrak{I}$  ist Modell von  $\text{REL}(x^1, x^2, r)$   
 und  
 $\mathfrak{I}$  ist Modell von  $\text{REL}(x^2, x^3, r)$   
 und  
 $\dots$   
 $\mathfrak{I}$  ist Modell von  $\text{REL}(x^{\alpha-1}, x^\alpha, r)$ .

$\mathfrak{I}$  ist Modell von  $\text{OP}(x, \beta \downarrow^1_r)$  gdw  $\mathfrak{I}$  ist Modell von  $\text{OP}(x, \beta \uparrow^1_{\text{inv}(r)})$ .

$\mathfrak{I}$  ist Modell von  $\text{OP}(x, \hat{\uparrow}^1_r)$  gdw  $\mathfrak{I}$  ist Modell von  $\text{REL}(t, x^1, r)$   
 und  
 $\mathfrak{I}$  ist Modell von  $\text{REL}(x^1, x^2, r)$   
 und  
 $\mathfrak{I}$  ist Modell von  $\text{REL}(x^2, x^3, r)$   
 und

$\mathfrak{S}$  ist Modell von  $REL(x^{\Omega-1}, x^{\Omega}, r)$ .

$\mathfrak{S}$  ist Modell von  $OP(x, \uparrow^t_r)$       gdw  $\mathfrak{S}$  ist Modell von  $OP(x, \uparrow^t_{inv(r)})$ .  
 $\mathfrak{S}$  ist Modell von  $OP(x, \uparrow^t_r)$       gdw  $\mathfrak{S}$  ist Modell von  $OP(x, \uparrow^t_r)$   
 oder  
 $\mathfrak{S}$  ist Modell von  $OP(x, \uparrow^t_{inv(r)})$ .  
 $\mathfrak{S}$  ist Modell von  $OP(x, \leftrightarrow^t_r)$       gdw  $\mathfrak{S}$  ist Modell von  $REL(x^i, t, r)$   
 und  
 $\mathfrak{S}$  ist Modell von  $REL(x^i, x, r)$ .

wobei:

$t_i$  sind Terme (außer Operatorterme)

$x^i$  sind Variablen für Terme (außer Operatorterme)

$\alpha, \beta$  sind die Nach- bzw. Vorfahrenstufen von  $t$  bezüglich  $r$

$\Omega$  ist die Nachfahrenstufe von  $t$  bezüglich  $r$  bis zu den Wurzeln

die Funktion "inv" angewendet auf eine Relationsart ergibt dessen inverse Relationsart

3.  $\mathfrak{S}$  ist Modell von  $\neg F$       gdw  $\mathfrak{S}$  ist nicht Modell von  $F$ .  
 $\mathfrak{S}$  ist Modell von  $(F_1 \wedge F_2)$       gdw  $\mathfrak{S}$  ist Modell von  $F_1$  und  $\mathfrak{S}$  ist Modell von  $F_2$ .  
 $\mathfrak{S}$  ist Modell von  $(F_1 \vee F_2)$       gdw  $\mathfrak{S}$  ist Modell von  $F_1$  oder  $\mathfrak{S}$  ist Modell von  $F_2$ .  
 $\mathfrak{S}$  ist Modell von  $(\exists x_{ij}) F$       gdw  $\mathfrak{S}^{ek}_{xij}$  ist Modell von  $F$  für mindestens ein  $e_k \in E_i$ .  
 $\mathfrak{S}$  ist Modell von  $(\forall x_{ij}) F$       gdw  $\mathfrak{S}^{ek}_{xij}$  ist Modell von  $F$  für alle  $e_k \in E_i$ .  
 wobei:  $i, j, k \in \mathbb{N}$

**Anfragen**

$\mathfrak{S}((\lambda \text{ REL})(\lambda x_{ij})F) = \{ \langle \delta, e_k \rangle : \mathfrak{S}_{\text{REL}}^{\delta \text{ ek}} x_{ij} \text{ ist Modell von } F \}$ .

wobei  $\delta$  ist Prädikat über  $\Delta \times R$

$e_k$  ist Individuum aus  $\Delta$

$\mathfrak{S}((\lambda \text{ REL})(\lambda y)F) = \{ \langle \delta, r_k \rangle : \mathfrak{S}_{\text{REL}}^{\delta \text{ rk}} y \text{ ist Modell von } F \}$ .

wobei  $\delta$  ist Prädikat über  $\Delta \times R$

$r_k$  ist Individuum aus  $R$

Beispielanfragen und deren Auswertung mit HyQL werden in Anhang A1 gezeigt.

Die folgenden Anfragen werden mit Erweiterungen von HyQL formulierbar sein:

- Anfragen nach Informationseinheiten mit Freitext (siehe auch Reiner 1994)  
Welche Informationseinheiten beginnen mit dem Zeichen 'fahr' und stehen in einer Partitionsbeziehung zu einer Informationseinheit, die die Zeichen 'rad' enthalten?
- Anfragen nach Relationsarten mit Typen  
Welche Relationsarten sind 'reflexiv', 'symmetrisch' und 'transitiv'?
- Anfragen nach mathematischen und semantischen Eigenschaften von Relationsarten  
Welche mathematischen Eigenschaften hat die Relationsart 'Unterbegriff'?  
Welche semantischen Eigenschaften hat die Relationsart 'Unterbegriff'?
- Anfragen nach Typen von Informationseinheiten  
Welchen Typ hat die Informationseinheit 'Fahrzeug'?
- Anfragen nach Merkmalsarten  
Welche Merkmalsarten hat die Informationseinheit 'PKW'?
- Anfragen nach Merkmalen  
Welches Merkmal hat die Informationseinheit 'Opel' zur Merkmalsart 'Hersteller'?
- komplexe Anfragen  
Welche Informationseinheiten vom terminologiespezifischen Typ BEGRIFF stehen in einer hierarchischen Relationsart, in einer reflexiven, symmetrischen und transitiven Relationsart oder in einer Relationsart, die 'Deutsche Über-

setzung' heißt, zur Informationseinheit vom terminologiespezifischen Typ BEGRIFF mit dem Wert 'PKW' in Beziehung?

Welche Informationseinheiten vom elementaren Typ BILD stehen mit den Informationseinheiten vom zusammengesetzten Typ BEWEGTES BILD in einer Partitionsbeziehung?

Welche aus Texte, Bildern und Tönen zusammengesetzte Informationseinheiten existieren?

## 6. Navigations- und Darstellungsmittel

Für die Praxis stellt sich heraus, daß kommandoorientierte Anfragemöglichkeiten alleine für den Informationssuchenden nicht ausreichen, um relevante Information effizient auffinden zu können (Kuhlen 1991, Duval, Oliivié 1992).

Es sind deshalb Verfahren wichtig, die die übersichtliche Darstellung von und die Navigation in terminologischen Beständen zum Gegenstand haben. Trotz fehlender oder nicht präsenster Begrifflichkeiten können dadurch Anfragen exakt formuliert und außerdem die Suche nach Information, ohne direkt danach gesucht zu haben, Serendipity-Effekt genannt (Kuhlen 1991 38), unterstützt werden. Das Ziel des Information Retrieval besteht nicht darin, eine Frage zu beantworten, sondern darin, sie zu stellen.

"(...) indem hier eine eigene Dynamik nicht nur von Seiten des Fragenden sondern auch von der des Textes zustande kommt (...). Diese Dynamik ist der des zwischenmenschlichen Dialogs vergleichbar, indem hier das System eine dem verobjektivierten Verständnis gemäße Antwort gibt."

(Capurro nach Konrad 1992 6)

Zur Steigerung der Übersichtlichkeit terminologischer Informationsbestände existieren unterschiedliche Ansätze. So sind folgende Darstellungsmittel (Felber, Budin 1989) zu unterscheiden:

## alphabetische Darstellungen

Benziner	UB: Elektro	Opel
UB: Pkw	UB: Fiat	UB: Pkw
OB: Fahrzeug	UB: Mercedes	OB: Fahrzeug
Diesel	UB: Opel	Personenkraftwagen
UB: Lkw	Familienwagen	BS: Pkw
UB: Pkw	OB: Pkw	Pkw
OB: Fahrzeug	Fiat	UB: Familienwagen
BF: Dieselmotorkraftfahrzeug	UB: Pkw	UB: Sportwagen
Dieselmotorkraftfahrzeug	OB: Fahrzeug	OB: Benziner
BS: Diesel	Lastkraftwagen	OB: Diesel
Elektro	BS: Lkw	OB: Elektro
UB: Lkw	Lkw	OB: Fiat
UB: Pkw	OB: Diesel	OB: Mercedes
OB: Fahrzeug	OB: Elektro	OB: Opel
BF: Elektrofahrzeug	OB: Mercedes	BF: Personenkraftwagen
Elektrofahrzeug	BF: Lastkraftwagen	Sportwagen
BS: Elektro	Mercedes	OB: Pkw
Fahrzeug	UB: Lkw	
UB: Benziner	UB: Pkw	
UB: Diesel	OB: Fahrzeug	

**Abb. 5:** alphabetische Darstellung eines Beispielthesaurus

## hierarchische Darstellungen

- systematische Darstellungen
- listenförmige Begriffspläne
- Dezimal-Klassifikationen

- Inhaltsverzeichnisse
- Bestandspläne
- Fachwerkpläne.

Beispiel:

```

Ernte Musik
  Vokalmusik
    Chorgesang
    Sologesang
Unterhaltungsmusik
  Vokalmusik
    Chorgesang
    Sologesang
  
```

Abb. 6: polyhierarchische Listendarstellung

### graphische Darstellungen

- Graphendarstellung
- rechteckige und kreisförmige (Pfeil)-Diagramme
- Baumdiagramme

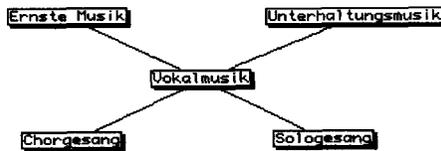


Abb. 7: Polyhierarchische Graphendarstellung nach DIN 1463:8.

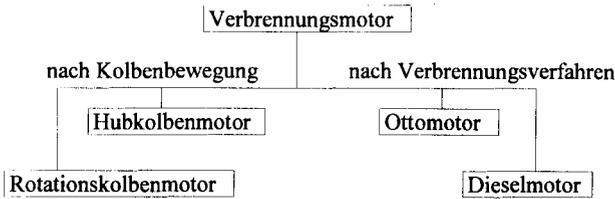


Abb 8: Polydimensionale Graphendarstellung nach DIN 1463:6.

## Pfade

## Glossare

## Inhaltsverzeichnisse, sachgebietsorientierte Darstellungen

### 2. Pflanze, Tier, Mensch (Körperliches)

- |                          |                       |
|--------------------------|-----------------------|
| 2.1. Pflanze             | 2.8. Tier             |
| 2.2. Pflanzenarten       | 2.9. Tierarten        |
| 2.3. Pflanzenteile       | 2.10. Tierzucht       |
| 2.4. Pflanzenkrankheiten | 2.11. Jagd            |
| 2.5. Pflanzenanbau       | 2.12. Tierkrankheiten |
| 2.6. Fruchtbarkeit       | 2.13. Mensch          |
| 2.7. Unfruchtbarkeit     | [...]                 |

Abb. 9: Sachgebietsorientierte Darstellung (nach Dornseiff 1959 8)

## Sortierung

- nach alphabetischer Reihenfolge
- nach Autor
- nach Entstehungszeitpunkt usw.

## Filter (Aus- und Einblenden von Information, Fischaugenprinzip)

### Beispiele:

alle Informationseinheiten, die von 'Müller' erzeugt wurden

alle Nachbartermini von 'Europa' bis zur Tiefe 2 und entlang der Teil-von-Relation

## (Mehrfach)-Markierungen (vgl. Rüter 1993)

### Beispiel:

Relationsarten werden fett und Definitionen werden kursiv angezeigt

## Navigationsmittel

Ein weiteres Mittel zur Verbesserung der Qualität und Effektivität des terminologiebasierten Information Retrieval sind Navigationswerkzeuge:

- Springen von Terminus-zu-Terminus entlang einer spezifizierten Relation  
Beispiel: von 'Fahrzeug' zu 'Auto' entlang der 'UB'-Relation
- die momentane Position im Netz anzeigen (Orten)  
Beispiel:
  - die Position in der alphabetischen Anordnung (der 2341'ste Knoten auf S. 20)
  - die Position in der räumlichen Anordnung  
(2. Hierarchiestufe zu 'Fahrzeug' entlang der 'UB'- Relation)
  - die Position im Inhaltsverzeichnis (Sachgebiet: Umwelt)
- die Umgebung der momentanen Position abtasten (Echolot)  
Beispiel: alle Nachbartermini bis zur Tiefe 2 vom Typ 'Multimedia'
- Markierung von Knoten
- Markierung von Wegen
  - geführte Wege in ausgesuchten Teilbereichen (guided tour)
  - vom Benutzer gegangene Wege (breadcrumbs)
  - Leser- und autorendefinierte Lesezeichen
- gleichzeitiges Beschreiten paralleler Wege  
Beispiel: von 'Fahrzeug' entlang aller 'UB'-Relationen
- Backtrack-Funktionen (schrittweise Zurücknahme durchgeführter Aktionen)  
Beispiel: den zuletzt besuchten Knoten aufsuchen
- Ausblenden von Teilbereichen  
Beispiel: nur die Pfade anzeigen, die den Namen 'Unterbegriff' haben
- Ikonische und symbolische Gestalten (Kuhlen 1991 111)

## 7. Ausblick

Für die Zukunft und für die Akzeptanz des terminologie- und hypertextbasierten Information Retrieval ist es wichtig, daß die technologischen mit den theoretisch formulierten Entwicklungen Schritt halten. Äußerungen, wie sie Bill Gates jüngst in einem Interview (Gates 1993) gab, er sprach von einer Datenautobahn für die USA, sollten die Wissenschaftler und Politiker auch in Deutschland und Europa dazu ermuntern, die technologische Infrastruktur für den möglichen qualitativen

Sprung auf ein hochwertigeres und effizienteres Informationsmanagement zu schaffen. Dafür sind die Anstrengungen auf dem Gebiet der großen und schnellen Datennetze (zum Beispiel mit Hilfe von ATM) und der objektorientierten Datenbanken (Ssykor 1994) voranzutreiben. Es gilt, die Flut von Informationsvermittlungsstellen massiv zu verringern, indem Information Retrieval Techniken aus dem Bereich Hypermedia (auch multilingual) eingesetzt werden. Öffentliche Bibliotheken sollten ihre Vorreiterrolle auf dem Gebiet des Informationsmanagement nicht noch weiter an die private Konkurrenz verspielen, sondern in ihren jeweiligen Spezialbereichen ihre Massenware, die Information, mit neuen Technologien öffentlich dem weltweit größtem Publikum, nämlich dem zu Hause, präsentieren. Die frei werdenden Mittel können die dann nötig werdenden Forschungsanstrengungen wie zum Beispiel die Qualitäts- und Konsistenzsicherung von Informationsbeständen oder die Entwicklung einheitlicher Anfragesprachen mitfinanzieren.

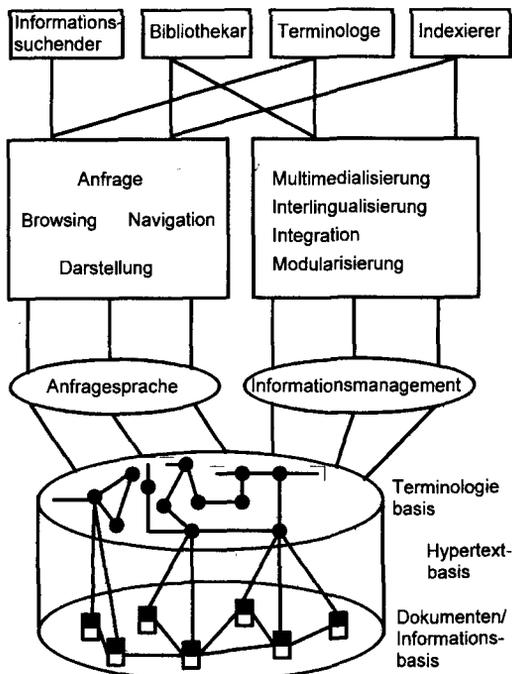


Abb. 10: Terminologie- und hypertextbasiertes Informationsmanagement.

## Anhang A1. Beispielanfragen mit der Sprache HyQL

### Informationseinheiten

$E_1 =$	{Fahrzeug, Auto, Opel, Mercedes, VW, Fiat}.	"Deskriptoren"
$E_2 =$	{Pkw, Karre, Wagen}.	"Nicht-Deskriptoren"
$E_3 =$	{"Fahrzeuge mit Benzinmotor stinken.", "Wer sein Auto liebt, der schiebt.", "Jeder Popel fährt nen Opel."}.	"Kommentare"
$E_4 =$	{vehicle, car, opel, mercedes, rabbit, fiat}.	"Englische Übersetzungen"
$E_5 =$	{DOC1, DOC2, DOC3}.	"Dokumente"
$\Delta =$	{ $E_1, E_2, E_3, E_4, E_5, E_1 \times E_1, E_1 \times E_2, E_1 \times E_3, \dots$ }.	

### Relationsarten

$R =$	{Unterbegriff, Oberbegriff, Synonym, Vorzugsbegriff, Teil_von, Be- steht_aus, Deutsch, Englisch, Kommentar, Indexiert_mit}
-------	---

### Beziehungen

REL(Fahrzeug, Auto, Unterbegriff), REL(Auto, Fahrzeug, Oberbegriff),  
 REL(Auto, Opel, Unterbegriff), REL(Opel, Auto, Oberbegriff),  
 REL(Auto, Mercedes, Unterbegriff), REL(Mercedes, Auto, Oberbegriff),  
 REL(Auto, VW, Unterbegriff), REL(VW, Auto, Oberbegriff),  
 REL(Auto, Fiat, Unterbegriff), REL(Fiat, Auto, Oberbegriff),

REL(Auto, Pkw, Synonym), REL(Pkw, Auto, Vorzugsbegriffs),  
 REL(Auto, Karre, Synonym), REL(Karre, Auto, Vorzugsbegriff),  
 REL(Auto, Wagen, Synonym), REL(Wagen, Auto, Vorzugsbegriff),

REL(Fahrzeug, "Fahrzeuge mit Benzinmotor stinken.", Kommentar),  
 REL(Auto, "Wer sein Auto liebt, der schiebt.", Kommentar),  
 REL(Opel, "Jeder Popel fährt nen Opel.", Kommentar),

REL(Fahrzeug, vehicle, Englisch), REL(vehicle, Fahrzeug, Deutsch),  
 REL(Auto, car, Englisch), REL(car, Auto, Deutsch),

REL(Opel, opel, Englisch), REL(opel, Opel, Deutsch),  
 REL(Mercedes, mercedes, Englisch), REL(mercedes, Mercedes, Deutsch),  
 REL(VW, rabbit, Englisch), REL(rabbit, VW, Deutsch),  
 REL(Fiat, fiat, Englisch), REL(fiat, Fiat, Deutsch),

REL(DOC1, Auto, Indexiert\_mit), REL(DOC1, Opel, Indexiert\_mit),  
 REL(DOC1, Mercedes, Indexiert\_mit), REL(DOC2, VW, Indexiert\_mit),  
 REL(DOC3, Fiat, Indexiert\_mit).

### Variablen

$x: \Delta; x_1: E_1; x_2: E_2; x_3: E_3; x_4: E_4; x_5: E_5$

$y: R$

**Frage 1:** Welche Informationseinheiten stehen zu Auto in einer Unterbegriffs-  
 beziehung?

$\mathfrak{S}((\lambda x) \text{REL}(\text{Auto}, x, \text{Unterbegriff})) =$

$\{\Delta: \mathfrak{S}_x^e \text{ ist Modell von } \text{REL}(\text{Auto}, x, \text{Unterbegriff})\} =$

$\{\Delta: \langle \mathfrak{S}_{x_i}^{el}(\text{Auto}), \mathfrak{S}_x^e(x), \mathfrak{S}_x^e(\text{Unterbegriff}) \rangle \in \mathfrak{S}_x^e(\text{REL})\} =$

$\{\Delta: \langle \text{Auto}, \Delta, \text{Unterbegriff} \rangle \in \mathfrak{S}(\text{REL})\} =$

$\{\text{Opel, Mercedes, VW, Fiat}\}$

**Frage 2:** Welche Relationsarten bestehen zwischen Auto und Opel?

$\mathfrak{S}((\lambda y) \text{REL}(\text{Auto}, \text{Opel}, y)) =$

$\{R: \mathfrak{S}_y^r \text{ ist Modell von } \text{REL}(\text{Auto}, \text{Opel}, y)\} =$

$\{R: \langle \mathfrak{S}_y^r(\text{Auto}), \mathfrak{S}_y^r(\text{Opel}), \mathfrak{S}_y^r(y) \rangle \in \mathfrak{S}_y^r(\text{REL})\} =$

$\{R: \langle \text{Auto}, \text{Opel}, R \rangle \in \mathfrak{S}(\text{REL})\} =$

$\{\text{Unterbegriff}\}$

**Frage 3:** Welche Relationsarten bestehen zwischen Auto und einer beliebigen  
 anderen Informationseinheit?

$$\mathfrak{S}((\lambda y) \text{REL}(\text{Auto}, x, y)) =$$

$$\{\text{R: } \mathfrak{S}_y^r \text{ ist Modell von } \text{REL}(\text{Auto}, x, y)\} =$$

$$\{\text{R: } \langle \mathfrak{S}_y^r(\text{Auto}), \mathfrak{S}_y^r(x), \mathfrak{S}_y^r(y) \rangle \in \mathfrak{S}_y^r(\text{REL})\} =$$

$$\{\text{R: } \langle \text{Auto}, \Delta, \text{R} \rangle \in \mathfrak{S}(\text{REL})\} =$$

$$\{\text{Unterbegriff, Oberbegriff, Synonym, Kommentar, Englisch}\}$$

**Frage 4:** Welche Beziehungen bestehen zwischen Auto und einer beliebigen anderen Informationseinheit mit einer beliebigen Relationsart?

$$\mathfrak{S}((\lambda \text{REL}) \text{REL}(\text{Auto}, x, y)) =$$

$$\{\delta: \mathfrak{S}_{\text{REL}}^\delta \text{ ist Modell von } \text{REL}(\text{Auto}, x, y)\} =$$

$$\{\text{REL}(\text{Auto}, \text{Fahrzeug}, \text{Oberbegriff}), \text{REL}(\text{Auto}, \text{Opel}, \text{Unterbegriff}),$$

$$\text{REL}(\text{Auto}, \text{Mercedes}, \text{Unterbegriffs-Relation}), \text{REL}(\text{Auto}, \text{VW}, \text{Unterbegriff}),$$

$$\text{REL}(\text{Auto}, \text{Fiat}, \text{Unterbegriff}), \text{REL}(\text{Auto}, \text{Pkw}, \text{Synonym}),$$

$$\text{REL}(\text{Auto}, \text{Karre}, \text{Synonym}), \text{REL}(\text{Auto}, \text{Wagen}, \text{Synonym}),$$

$$\text{REL}(\text{Auto}, \text{"Wer sein Auto liebt, der schiebt."}, \text{Kommentar}),$$

$$\text{REL}(\text{Auto}, \text{car}, \text{Englisch})\}$$

**Frage 5:** Welches sind die Unterbegriffe von Fahrzeug?

$$\mathfrak{S}((\lambda x) \text{OP}(x, \uparrow_{\text{Unterbegriff}}^{\text{Fahrzeug}})) =$$

$$\{\Delta: \mathfrak{S}_x^e \text{ ist Modell von } \text{OP}(x, \uparrow_{\text{Unterbegriff}}^{\text{Fahrzeug}})\} =$$

$$\{\Delta: \mathfrak{S}_x^e \text{ ist Modell von } \text{REL}(\text{Fahrzeug}, x^1, \text{Unterbegriff}) \text{ und}$$

$$\mathfrak{S}_x^e \text{ ist Modell von } \text{REL}(x^1, x^2, \text{Unterbegriff})\} =$$

$$\{\Delta: \langle \mathfrak{S}_x^e(\text{Fahrzeug}), \mathfrak{S}_x^e(x^1), \mathfrak{S}_x^e(\text{Unterbegriff}) \rangle \in \mathfrak{S}_x^e(\text{REL}) \text{ und}$$

$$\langle \mathfrak{S}_x^e(x^1), \mathfrak{S}_x^e(x^2), \mathfrak{S}_x^e(\text{Unterbegriff}) \rangle \in \mathfrak{S}_x^e(\text{REL})\} =$$

$$\{\text{Auto, Opel, Mercedes, VW, Fiat}\}$$

**Frage 6:** Welche Dokumente sind mit Fahrzeug oder einem seiner Unterbegriffe indexiert?

$$\mathfrak{S}((\lambda x_5) \text{REL}(x_5, x, \text{Indexiert\_mit}) \wedge \text{OP}(x, \uparrow_{\text{Unterbegriff}}^{\text{Fahrzeug}})) =$$

$$\{\text{E}_5: \mathfrak{S}_{x_5}^{\text{E}_5} \text{ ist Modell von } \text{REL}(x_5, x, \text{Indexiert\_mit}) \wedge \text{OP}(x, \uparrow_{\text{Unterbegriff}}^{\text{Fahrzeug}})\} =$$

$$\{\text{E}_5: \mathfrak{S}_{x_5}^{\text{E}_5} \text{ ist Modell von } \text{REL}(x_5, x, \text{Indexiert\_mit}) \text{ und}$$

$$\mathfrak{S}_{x_5}^{\text{E}_5} \text{ ist Modell von } \text{OP}(x, \uparrow_{\text{Unterbegriff}}^{\text{Fahrzeug}})\} =$$

$\{E_5: \langle \mathfrak{S}_{x_5}^{e_5}(x_5), \mathfrak{S}_{x_5}^{e_5}(x), \mathfrak{S}_{x_5}^{e_5}(\text{Indexiert\_mit}) \rangle \in \mathfrak{S}_{x_5}^{e_5}(\text{REL}) \text{ und}$   
 $\mathfrak{S}_{x_5}^{e_5} \text{ ist Modell von REL(Fahrzeug, } x^1, \text{ Unterbegriff) und}$   
 $\mathfrak{S}_{x_5}^{e_5} \text{ ist Modell von REL}(x^1, x^2, \text{ Unterbegriff))} =$   
 $\{E_5: \langle \mathfrak{S}_{x_5}^{e_5}(x_5), \mathfrak{S}_{x_5}^{e_5}(x^1), \mathfrak{S}_{x_5}^{e_5}(\text{Indexiert\_mit}) \rangle \in \mathfrak{S}_{x_5}^{e_5}(\text{REL}) \text{ und}$   
 $\langle \mathfrak{S}_{x_5}^{e_5}(\text{Fahrzeug}), \mathfrak{S}_{x_5}^{e_5}(x^1), \mathfrak{S}_{x_5}^{e_5}(\text{Unterbegriff}) \rangle \in \mathfrak{S}_{x_5}^{e_5}(\text{REL}) \text{ und}$   
 $\langle \mathfrak{S}_{x_5}^{e_5}(x^1), \mathfrak{S}_{x_5}^{e_5}(x^2), \mathfrak{S}_{x_5}^{e_5}(\text{Unterbegriff}) \rangle \in \mathfrak{S}_{x_5}^{e_5}(\text{REL})\} =$   
 { DOC1, DOC2, DOC3}

## Anhang A2. Normen

### National:

**DIN 1463 Teil 1:** (Nov. (1987) Erstellung und Weiterentwicklung von Thesauri; Einsprachige Thesauri

**DIN 1463 Teil 2:** (Entwurf Dez. 1988) Erstellung und Weiterentwicklung von Thesauri; Mehrsprachige Thesauri

**DIN 2330:** (März 1979) Begriffe und Benennungen; Allgemeine Grundsätze

**DIN 2331:** Begriffssysteme und ihre Darstellung

**DIN 2332:** Benennen international übereinstimmender Begriffe

**DIN 2333:** (Dez. 1987) Fachwörterbücher, Stufen der Ausarbeitung

**DIN 2335:** (Okt. 1986) Sprachenzeichen

**DIN 2336:** (März 1979) Lexikographische Zeichen für manuell erstellte Fachwörterbücher

**DIN 2339 Teil 1:** (Mai 1987) Ausarbeitung und Gestaltung von Veröffentlichungen mit terminologischen Festlegungen

**DIN 2339 Teil 2:** (Entwurf August 1986) Ausarbeitung und Gestaltung von Veröffentlichungen mit terminologischen Festlegungen

**DIN 2341 Teil 1:** (Entwurf Okt. 1986) Format für den maschinellen Austausch terminologischer/lexikographischer Daten - MATER; Kategorienkatalog

**DIN 2342 Teil 1:** (Entwurf Dez. 1986) Begriffe der Terminologielehre

**DIN 31623 Teil 1 - 3:** (September 1988) Indexierung zur inhaltlichen Erschließung von Dokumenten

**RAK.** Regeln für die alphabetische Katalogisierung. Wiesbaden, Reichert, 1977.

**RSWK.** Regeln für den Schlagwortkatalog. RSWK/bearbeitet von der Kommission des DBI für Sacherschließung. Berlin, Deutsches Bibliotheksinstitut, 1991.

**Schlagwortnormdatei.** Schlagwortnormdatei (SWD). bearb. von der Deutschen Bibliothek, Frankfurt am Main, Deutsche Bibliothek, 1993.

International:

**ISO 704.** Principles and methods of terminology. Genf, ISO, 1987.

**ISO 1087.** Terminology - Vocabulary. Genf, ISO, 1990.

**ISO 1087-2 (Committee Draft).** Terminology work - Vocabulary - Part 2: Computational aids in terminology. Genf, ISO/TC 37/SC 3/N 127, 1993.

**ISO 1951.** Lexicographical symbols particularly for use in classified defining vocabularies. Genf, ISO, 1973.

**ISO 2709.** Documentation - Format for bibliographic information interchange on magnetic tape. Genf, ISO, 1981.

**ISO 2788.** Documentation - Guidelines for the establishment and development of monolingual thesauri. Genf, ISO, 1986.

**ISO 5127-2.** Documentation and information - Vocabulary - Part 2 : Traditional documents. Genf, ISO, 1983.

**ISO 5127-3A.** Documentation and information - Vocabulary - Section 3a) : Acquisition, identification, and analysis of documents and data. Genf, ISO, 1981.

**ISO 5127-6.** Documentation and information - Vocabulary - Part 6 : Documentary languages. Genf, ISO, 1983.

**ISO 5127-11.** Documentation and information - Vocabulary - Part 11 : Audio-visual documents. Genf, ISO, 1987.

**ISO 5963.** Documentation - Methods for examining documents, determining their subjects, and selecting indexing terms. Genf, ISO, 1985.

**ISO 5964.** Documentation - Guidelines for the establishment and development of multilingual thesauri. Genf, ISO, 1985.

**ISO 6156.** Magnetic tape exchange format for terminological/lexicographical records (MATER). Genf, ISO, 1987 (E).

**ISO 8879.** Information Processing - Text and office systems - Standard Generalized Mark-up Language (SGML). Genf, ISO, 1986.

**ISO 10744.** Information Technology - Hypermedia/Time-based Structuring Language (HyTime). Genf, ISO, 1992.

**ISO 12200 (Committee Draft).** Computational aids in terminology - Terminological interchange format (TIF) - An SGML application. Genf, ISO/TC 37/SC 3/WG 3/N9, 1993.

**ISO 12620 (Working Draft).** Computational aids in terminology - Terminology interchange format (TIF) - Data element dictionary. Genf, ISO/TC 37/SC 3/WG 1/N 7, 1993.

## **Internationale Arbeitsgruppen für Standardisierung:**

Technical Committee ISO TC 37: Terminology

Technical Committee ISO TC 46: Documentation

Joint Technical Committee ISO/IEC JTC 1: Information Technology

Einen Überblick über die offene Bearbeitung multimedialer Dokumente und diesbezüglicher Normungsprojekte gibt Bormann, Bormann 1991. Europaweit verfügbare Thesaurusbestände (knapp 1000) werden in Thesaurus Guide 1985 aufgeführt. Einen ausführlichen Vergleich von Thesaurussoftware anhand eines ausgewählten Kriterienkataloges gibt Schönfeld 1994. Weltweites Koordinationszentrum für terminologische Fragestellungen aller Art ist TermNet in Wien (Felber, Budin 1989).

## **Literatur**

- Afrati, Koutras (1990):** Afrati, F. und C.D. Koutras. A hypertext model supporting query mechanisms. In André, Rizk, Streitz (1990): 52-66.
- Beeri, Kornatzky (1990):** Beeri, C. und Y. Kornatzky. A logical query language for hypertext systems. In: André, Rizk, Streitz (1990): 67-80.
- Böhm, Mengel, Muhr (1994):** Böhm, Andreas, Andreas Mengel und Thomas Muhr (Hrsg.). Texte verstehen - Konzepte, Werkzeuge und Methoden. Schriftenreihe zur Informationswissenschaft, Universitätsverlag Konstanz.
- Bormann, Bormann (1991):** Bormann, Ute; Bormann Carsten. Offene Bearbeitung multimedialer Dokumente. In: Informatik Spektrum, 14, S. 270-280. Springer Verlag, 1991.
- Conklin (1987):** Conklin, J. Hypertext - An introduction and a survey. In: IEEE Computer 20 9, S. 17-41, 1987.
- Czedik (1992):** Czedik, Dorothea. Status Quo der Wiederverwendbarkeit von Wissensbasen. In: Künstliche Intelligenz, Zeitschrift der Gesellschaft für Informatik e.V. 27-32, März 1992.
- Dornseiff (1959):** Dornseiff, Franz. Der Deutsche Wortschatz nach Sachgruppen. Berlin, De Gruyter, 1959.
- Duval, Olivie (1992):** Duval, E. und H. Olivie. Towards the Integration of a Query Mechanism and Navigation for Retrieval of Data on Multimedia Documents. In: ACM SIGIR FORUM Vol 26, Nr.2, S. 8, July 1992.
- Felber, Budin (1989):** Felber, Helmut und Gerhard Budin. Terminologie in Theorie und Praxis. Tübingen, Narr, 1989.
- Fischer, Möhr (1991):** Fischer, Dietrich; Wiebke Möhr. Lexikonredaktion: eine Herausforderung für Computerassistenten beim Publizieren. In: GMD-Spiegel, 1/1991.
- Fowler, Fowler, Wilson (1991):** Fowler, Richard, Wendy Fowler und Bradley Wilson. Integrating Query, Thesaurus, and Documents through a Common Visual Representation. In: Proceedings of the International ACM SIGIR Conference on Research and Development in Information Retrieval, 142-152, 1991.

- Fuhr (1991):** Fuhr, Norbert (Hrsg.). Information Retrieval. Informatik Fachberichte Nr. 289. Berlin, Springer Verlag, 1991.
- Gates (1993):** Gates, Bill. Wir bauen die Datenautobahn. In: Der Spiegel, Heft 20, 1993.
- Konrad, Reiner (1986):** Konrad, Erhard; Ulrike Reiner. Zur Semantik von Anfragesprachen für Dokumentennachweissysteme. In: Tagungsband Deutscher Dokumentartag 1985, Nürnberg, Oktober 1985. München, Saur, 180-197, 1986.
- Konrad (1992):** Konrad, Erhard. Zur Effektivitätsbewertung von Information-Retrieval-Systemen. In: Experimentielles und Praktisches Information Retrieval (Hrsg. Kuhlen). Schriften zur Informationswissenschaft Bd.III. 119-130. Universitätsbibliothek Konstanz 1992.
- Kuhlen (1991):** Kuhlen, Rainer. Hypertext, ein nicht lineares Medium zwischen Buch und Wissensbank. Springer Verlag, Berlin, 1991.
- McMath, Rada, Tamaru (1989):** McMath, C.F., R. Rada und R.S. Tamaru. A graphical thesaurus-based information retrieval system. In: International Journal of Man Machine Studies, 31, 121-147, 1989.
- Mengel (1991):** Mengel, Andreas. Thesaurusrelationen, Konsistenz, Inferenz und Interdependenz. Forschungsbericht Nr. 91-5, Projekt ATLAS, TU Berlin 1991.
- Neubauer (1993):** Neubauer, Wolfram (Hrsg.). Deutscher Dokumentartag 1992. Proceedings. Deutsche Gesellschaft für Dokumentation, Frankfurt am Main, 1993.
- Pollard (1990):** Pollard, Richard. Hypertext presentation of thesauri used in online searching. In: Electronic Publishing, Vol. 3(3), S. 155-172, August 1990.
- Reiner (1991):** Reiner, Ulrike. Anfragesprachen für Informationssysteme. Reihe Informationswissenschaft der DGD, Bd.I. Frankfurt a.Main, Deutsche Gesellschaft für Dokumentation, 1991.
- Reiner (1994):** Reiner, Ulrike. Anfragesprachen für Textsuchsysteme. In: Böhm, Mengel, Muhr 1994.
- Rizk, Streitz, André (1990):** Rizk, A., Norbert Streitz und J. André. (Hrsg.). Hypertext: Concepts, Systems and Applications. Cambridge: University Press, 1990.
- Rüster (1993):** Rüster, Armin. Hypertext-Konzepte für die Navigation in terminologisch/lexikographischen Wortbeständen. Unveröff. Diplomarbeit an der TU Berlin, Institut für Angewandte Informatik, 1993.
- Salton, McGill (1987):** Salton, Gerard and Michael McGill. Information Retrieval - Grundlegendes für Informationswissenschaftler. Hamburg, McGraw-Hill 1987.
- Schönfeldt (1993):** Schönfeldt, René. Terminologisch/lexikographische Wortsammlungen. Unveröff. Studienarbeit an der TU Berlin, Institut für Angewandte Informatik, 1993.
- Ssykor (1994):** Mathias Ssykor. Kopplung der Programmiersprache Smalltalk-80 mit der objektorientierten Datenbank GemStone für die Thesauruspflege. Unveröff. Studienarbeit an der TU Berlin, Institut für Angewandte Informatik, 1994.
- Thesaurus Guide (1985):** Thesaurus Guide. Luxembourg, Office for Official Publications of the European Communities, 1985.
- Willenborg (1991):** Willenborg, Josef. ATLAS-PfleSaurus: Ein objektorientiertes System zur Unterstützung der Thesauruspflege. In: Fuhr 1991, 51-63.
- Willenborg (1993):** Willenborg, Josef. Hypermediabasierte Terminologie- und Wörterbuchpflege. In: Neubauer 1993.
- Willenborg (1994):** Willenborg, Josef. PflerSaurus - ein Werkzeug zur Thesauruspflege. In: Böhm, Mengel, Muhr 1994.