

**State of the Art auf dem Gebiet der
Repräsentation und Verarbeitung
von Thesauri am Beispiel der Erstellung
eines Thesaurus für eine
Firmenbibliothek**

Diplomarbeit

im Fach

der

Fachhochschule Stuttgart –

Hochschule der Medien

Reinhard Müller

Erstprüfer:

Prof. Dr. Wolf-Fritz Riekert

Zweitprüferin:

Prof. Margarete Payer

Bearbeitungszeitraum: 07. 06 2003 bis 07. 10. 2003

Stuttgart, September 2003

Kurzfassung

In dieser Diplomarbeit wird nach einer allgemeinen Einführung und dem Vergleich zwischen dem Umweltthesaurus des Umweltbundesamtes und General Multilingual Environmental Thesaurus (GEMET) ein Thesaurus mit vorgegebenen Schlagwörtern aus der Firmendatenbank für eine lokal ansässige Firma entworfen. Dieser wird mit Thesauruswerkzeugen, im Besonderen mit dem Programm „thew 32“, verwirklicht. Ferner wird die Möglichkeit, Thesauri mit XML zu verwirklichen, untersucht.

Schlagwörter: Thesaurus, Thesauri, state of the art, Entwicklung, lokale Firma

Abstract

In this diploma thesis a thesaurus is developed for a local firm. After a general introduction and a comparison of the environmental thesaurus of the Umweltbundesamt (German ministry of environment) and General Multilingual Environmental Thesaurus (GEMET), it is made by using a tool for thesauri called “thew 32” and using keywords provided by the firms data base. Furthermore, it is pointed out to do thesauri with XML.

Key word: Thesaurus, Thesauri, state of the art, development, local firm

Inhaltsverzeichnis

Kurzfassung	2
Inhaltsverzeichnis	3
Abbildungsverzeichnis	5
Tabellenverzeichnis	5
Abkürzungsverzeichnis	6
Vorwort	7
1 Überblick	8
2 Einführung	9
2.1 Ausgangslage.....	9
2.2 Herangehensweise an die Aufgabenstellung	9
3 Definition und Historie des Thesaurusbegriffs	11
3.1 Arten von Thesauri	11
3.1.1 Semiotische Thesauri.....	12
3.2 Abgrenzung des Thesaurus zu anderen Dokumentations Sprachen.....	12
3.3 Verschiedene Thesaurusdefinitionen.....	15
3.4 Verschiedene Relationstypen	15
3.4.1 Äquivalenzrelation	16
3.4.2 Generische Relation.....	17
3.4.3 Partitive Relation	17
3.4.4 Assoziationsrelation	17
3.4.5 Begriffskombination.....	18
3.5 Probleme bei der Thesauruserstellung.....	18
3.5.1 Synonymproblem	19
3.5.2 Polysemproblem	20
3.5.3 Deskriptoren	20
3.5.4 Non - Deskriptoren	20
3.6 Arten von Thesaurussoftware.....	20
3.6.1 Eigenständige Thesaurussoftware	21
3.6.2 Integrierte Thesaurussoftware	21
3.6.3 Anforderungen an eine Thesaurussoftware	21
4. Verschiedene Umweltthesauri	22

4.1	Umweltthesaurus des Umweltbundesamtes	22
4.1.2	Aufgaben des Umweltbundesamtes Berlin	22
4.1.3	Umweltdatenbanken des Umweltbundesamtes	22
4.1.4	Der Umweltthesaurus des Umweltbundesamtes	23
4.1.5	Aufgaben des Umweltbundesamtes auf nationaler und internationaler Ebene ...	23
4.2	General Multilingual Environmental Thesaurus (GEMET).....	23
5.	Eingeschlagener Realisierungsweg	26
5.1	Import der Schlagworte und verschiedene Ordnungsarbeiten.....	26
5.2.	Vergleich zwischen GEMET und Bibliotheksschlagworten.....	30
5.3	Importieren der Schlagworte in Thesaurussoftware „Thew32“ und gegenseitige Verlinkung	31
5.4	Abschließende Arbeiten am Thesaurus	36
5.4.1	Umwandlung importierter Begriffe in thesauruseigene	36
5.4.2	Löschen der importierten Begriffe und Sonderzeichenauflösung.....	37
6.	Umwandlung der Thesaurusdatei in XML	39
7.	Zusammenfassung und Ausblick	40
Glossar		41
Quellenangaben.....		42
Erklärung		45

Abbildungsverzeichnis

Abbildung 1 Beispiel einer analytischen Klassifikation.....	13
Abbildung 2 Beispiel einer facettierten Klassifikation.....	13
Abbildung 3 Beispiele für Thesauri, welche mit „Thew 32“ realisiert wurden.....	26
Abbildung 4 Importierte Schlagwortliste aus Bibliotheksdatenbank	27
Abbildung 5 Entfernen von Duplikaten und Pluralformen.....	28
Abbildung 6 Entgeltige Schlagwortliste.....	29
Abbildung 7 Korrigierte GEMET Schlagwortliste.....	29
Abbildung 8 JOIN Abfrage mit dem Assistenten realisiert.....	30
Abbildung 9 Vergleich von GEMET (links) und Bibliotheksschlagworten (rechts)	30
Abbildung 10 Verlinkung der Hauptgruppen des Thesaurus	31
Abbildung 11 Verlinkung von Abkürzungen mit ihren Langformen.....	32
Abbildung 12 Verlinkung der Hauptbegriffe als „Broader Terms“.....	33
Abbildung 13 Schlagwortsuche im GEMET	33
Abbildung 14 Auswählen der Schlagwörter, die nicht im GEMET vorkommen.....	34
Abbildung 15 Fertige Liste der zu GEMET - nichtkonformen Schlagwörter	35
Abbildung 16 Ergänzen durch erweiterte Begriffe aus dem Thesaurus	36
Abbildung 17 Verwandlung importierter Begriffe in thesauruseigene.....	37
Abbildung 18 Das “modify reference/note“ – Fenster	37
Abbildung 19 Bestätigungsnachricht zur Löschung des Begriffs „Fauna“	38
Abbildung 20 fertiggestellter Thesaurus.....	38
Abbildung 21 Eingabemaske zur Umwandlung der Thesaurusdatei in XML.....	39
Abbildung 22 Ergebnis der Umwandlung in XML Format.....	39

Tabellenverzeichnis

Tabelle 1 DIN 1462 gibt für die Beziehungen zwischen den einzelnen Begriffen folgende Kürzel an:	15
Tabelle 2 ISO 2788 nennt folgende Begriffe:.....	16
Tabelle 3 Umweltdatenbanken des Umweltbundesamtes	22
Tabelle 4 Eine Zählung der Schlagworthäufigkeiten ergibt folgendes Resultat:	31

Abkürzungsverzeichnis

ADIS	adaptable Documentation and Information System
DIN	Deutsches Institut für Normung
EEA	European Environment Agency
ESI	Environnement Scientifique Intégré
ETC/CDS	European Topic Centre on Catalogue of Data Resources
HdM	Hochschule der Medien
ISO	International Organization for Standardization
IuD	Informations-, und Dokumentationswesen
PDF	Portable Document Format
SQL	Structured Query Language
XML	extensible Markup Language

Vorwort

Auf der Suche nach einem Diplomarbeitsthema im Studiengang Informationswirtschaft an der HdM, kam es mir in den Sinn, einen Thesaurus für die Bibliothek der Firma zu erstellen, für die ich nebenberuflich arbeite. Ich wollte so beiden Seiten, mir als Diplomanden und meiner Arbeitsstelle, durch eine Verbesserung der derzeitigen Lage, helfen. Deshalb wird in dieser Diplomarbeit, nach der Definition des Begriffes Thesaurus und eingehender Diskussion, ein Thesaurus für eine Firmenbibliothek realisiert.

Meinen Dank aussprechen möchte ich meinen beiden Prüfern, Herrn Prof. Dr. Wolf-Fritz Riekert und Frau Prof. Margarete Payer, die mir mit Rat und Tat zur Seite standen. Ebenso möchte ich Herrn Dr. Ing. Michael Koch, Inhaber des Planungsbüros „Planung und Umwelt“ und all seinen Mitarbeitern danken, die mich bei Fachfragen ebenso unterstützten.

Möge es dieser Arbeit gelingen, einen Einblick in die Konzeption und Realisierung eines Thesaurus zu verschaffen.

1 Überblick

Mit dieser Diplomarbeit wird ein Einblick in die Konzeption und Realisierung eines Umweltthesaurus für eine Firmenbibliothek gegeben.

Im nun folgenden Kapitel werden die Gegebenheiten, wie sie in der Firmenbibliothek vorlagen, die Aufgabenstellung der Arbeit und ihre Herangehensweise beschrieben.

Im nächsten Kapitel wird der Begriff Thesaurus etymologisch und historisch erklärt, definiert und anschließend gegen andere Begriffe abgegrenzt. Ferner werden spezielle Probleme bei der Thesauruserstellung erwähnt.

Im nachfolgenden Kapitel wird auf den derzeitigen Stand der Forschung eingegangen, insbesondere auf die Umweltthesauri des Umweltbundesamtes und den europäischen General Multilingual Environmental Thesaurus (GEMET).

Im darauffolgenden Kapitel wird auf die Realisierung eines Thesaurus, mit der Software "thew 32" (Craven, 2003) eingegangen.

Im letzten Kapitel wird eine Zusammenfassung und ein Ausblick auf fortführende Arbeiten gegeben.

Englischsprachige Quellen werden vom Autor übersetzt.

2 Einführung

In dieser Einführung wird das Vorgehen bei der Erarbeitung des Firmenthesaurus geschildert.

2.1 Ausgangslage

Die Firmenbibliothek, an welcher der Verfasser im Nebenjob arbeitet, hat rund 3000 Medien, vornehmlich Bücher, aber auch CD-ROMS, die aber bei der Katalogisierung wie Bücher behandelt werden. Zeitschriften werden nicht in den Katalog aufgenommen. Das Bibliothekssoftwareprogramm ist eine Eigenproduktion auf SQL Basis.

Da die alte Datenbank kaum benutzt wurde, habe ich in meiner Grossen Studienarbeit (Müller, 2003) ihre Daten in eine ACCESS Datenbank überführt. Hiervon sind noch Schlagwörter vorhanden. Diese sind alphabetisch geordnet und können den einzelnen Medien zugeordnet werden, bei denen sie dann in Schlagwortketten auftreten. Leider ist die Aufnahme neuer Schlagwörter begrenzt und es ist zweckmäßig, die einzelnen Schlagwörter nach Themengebieten geordnet und untereinander verlinkt, darzustellen. Darum habe ich es mir zur Aufgabe gemacht, einen Thesaurus zu entwerfen.

2.2 Herangehensweise an die Aufgabenstellung

Zunächst werden die Schlagworte aus der Datenbank exportiert, auf Fehler (wie doppeltes Vorhandensein, Rechtschreibfehler) überprüft, mit einer ausgedruckten alphabetischen Schlagwortliste verglichen, die noch einige Neuerungen enthalten kann und dann mit dem GEMET auf Wortgleichheit hin untersucht. Dies geschieht durch eine Abfrage in ACCESS.

Schließlich werden die Schlagworte nach dem GEMET Prinzip aufgebaut und wo nötig, durch Begriffe aus demselben ergänzt.

Zunächst muss entschieden werden, wie der Thesaurus aufgebaut sein soll. Aitchison und Gilchrist (1987, S. 5-6) diskutieren dieses Thema ebenfalls: welche Dinge (z.B. Periodika oder Bücher) werden im Thesaurus abgebildet? In welche Tiefe? Soll der Thesaurus Mono-, oder Multilingual sein? Wer sind die Endbenutzer? Soll es ein Such- oder Indexierungsthesaurus sein?

Dies Fragen beantworten sich in unserem Falle folgendermaßen: Der Thesaurus soll Bücher und andere einfache Medien wie CD-ROMS aber keine periodischen Zeitschriften enthalten. Der Thesaurus soll eher einfach gehalten sein, damit die Suche nicht zu aufwendig wird. Es wird sich um einen einsprachigen Thesaurus handeln, da die Be-

griffe sowohl der bisherigen Bibliotheksdatenbank, als auch des GEMET auf Deutsch vorliegen. Später können ja noch jederzeit sowohl neue Begriffe, als auch fremdsprachige Begriffe hinzugefügt werden. Als Endbenutzer kommt das Personal der Firmenbibliothek in Frage, auch deshalb sollte der Thesaurus nicht nur in seinem Aufbau, sondern auch in seiner Bedienung einfach sein.

Beim Thesaurus wird es sich um einen Suchthesaurus handeln, in dem Begriffe mit dem Suchfeld von „Thew 32“ oder durch durchklicken durch die Hierarchie gesucht werden können.

3 Definition und Historie des Thesaurusbegriffs

Der Begriff Thesaurus kommt aus dem Griechischen und bedeutet übersetzt „Schatzhaus“. Hierbei handelte es sich um seit dem 7. Jahrhundert vor Christus durch Fürsten und Städte erbaute Gebäude. Diese wurden bei Heiligtümern erbaut und enthielten wertvolle Weihegaben. Besonders bekannt sind die „Thesauroi“ von Olympia, Delos und Delphi (Lexikon, 1968). Als „Thesaurus“ wurden sowohl die Weihegaben als auch deren Behältnis bezeichnet (Laisiepen et al. 1972 S. 271)

1265 bezeichnete der Begriff erstmals eine systematisierende Enzyklopädie, durch den Florentiner Brunetto Latini, der sein Werk „Li Livres dou Tresor“ nannte (Umstätter, 1999 S. 3).

Später wurde der Begriff des Wertvollen dann auf alphabetisch oder thematisch geordnete Wörtersammlungen besonders antiker Sprachen übertragen. Sie sollten eine Sprache in Ihrer Gesamtheit erschließen. (Reusch, 2001)

Beispiele sind hier der „Thesaurus Linguae Graecae“ des Griechischen und „Thesaurus Linguae Latinae“ für das Lateinische (Lexikon, 1968).

Peter Mark Roget, ein englischer Arzt, der 1852 einen Thesaurus herausgab, in dem neben Synonymen auch die zu einem Wort begrifflich dazugehörenden Wörter angegeben wurden, lieferte die Version, an welche heute das Informations- und Dokumentationswesen anknüpft.

Seit 1965 gehört der Thesaurus zum festen Inventar im Informations- und Dokumentationswesen.

Allerdings wurde der Begriff sehr früh schon auf verschiedene Gebiete angewendet, so dass eine Abgrenzung schwierig ist (Wersig 1978, S. 23).

3.1 Arten von Thesauri

„In der Lexikographie gibt es Definitionsthesauri, die eingetragene Termini erklären, während andere sie übersetzen. Wieder andere führen zu einem Term Synonyme und Antonyme auf und verweisen gegebenenfalls auf den dazugehörigen Wortschatz. Es gibt auch Mischformen zwischen den Typen.

Hingegen wird in der Dokumentationswissenschaft unter Thesauri eine thematische Sammlung von Wörtern eines Fachgebietes, zum Beispiel für Medizin, Chemie oder Biologie bezeichnet. In ihnen

werden Beziehungen eines jeden Wortes, nicht nur auf sprachlicher Ebene, aufgezeigt.“ (Reusch, 2001).

Mit letzteren Thesauri möchte ich mich im Folgenden beschäftigen.

3.1.1 Semiotische Thesauri

Umstätter (1999, S. 9) weist noch auf eine Neuentwicklung hin. In semiotischen Thesauri

„(erhalten) Worte und andere Bezeichnungen durch ihre Syntax und dabei insbesondere durch ihre hierarchische Organisation eine Begrifflichkeit, die auch für Computer verständlich ist (...) Diese können dann wieder den Endnutzern sinnvolle Antworten auf Ihre Fragen geben.“

Er veröffentlichte auf seiner Site einen semiotischen Thesaurus der Bibliothekswissenschaft ([Umstätter, 2003](#)).

3.2 Abgrenzung des Thesaurus zu anderen Dokumentationsprachen

Neben Thesauri gibt es hierarchische und facettierte Klassifikationen.

In ersteren, auch analytische Klassifikationen genannt,

„werden die in der Systematik zusammengeführten Begriffe entsprechend den Gegebenheiten des Wissensgebietes von oben nach unten, von Allgemeinen zum Speziellen, immer feiner untergliedert. Bei der Vergabe von Indizes darf nur, was in der Klassifikation präkoordiniert¹ enthalten ist, verwendet werden“ (Manecke, 1997 S. 145).

Bei facettierten oder analytisch-synthetischen Klassifikationen werden aus einer Systematik zusammengestellte, gleichrangige Merkmalsbegriffe eines Wissensgebietes, entsprechende Einzelbegriffe zugeordnet, die dann wieder weiter untergliedert werden können.

„Sie sind ahierarchisch und mehrdimensional. Mit ihnen können postkoordinativ (...) auch sehr komplexe Sachverhalte, (...) durch Zusammenführung (Synthese) von Begriffen aus verschiedenen Facetten (Begriffsgruppen) wiedergegeben werden“ (Manecke, 1997 S. 145).

Sie sind flexibler als hierarchische Klassifikationen und können zur Thesauruserstellung verwendet werden, indem sie ein Werkzeug zur Analyse des Wissensgebietes und

¹ Fremdwörter werden im Glossar erklärt

seiner Beziehungen sind. Auch können sie als eine systematische Anzeige für den fertigen Thesaurus dienen (Aitchison/Gilchrist, 1987 S. 50).

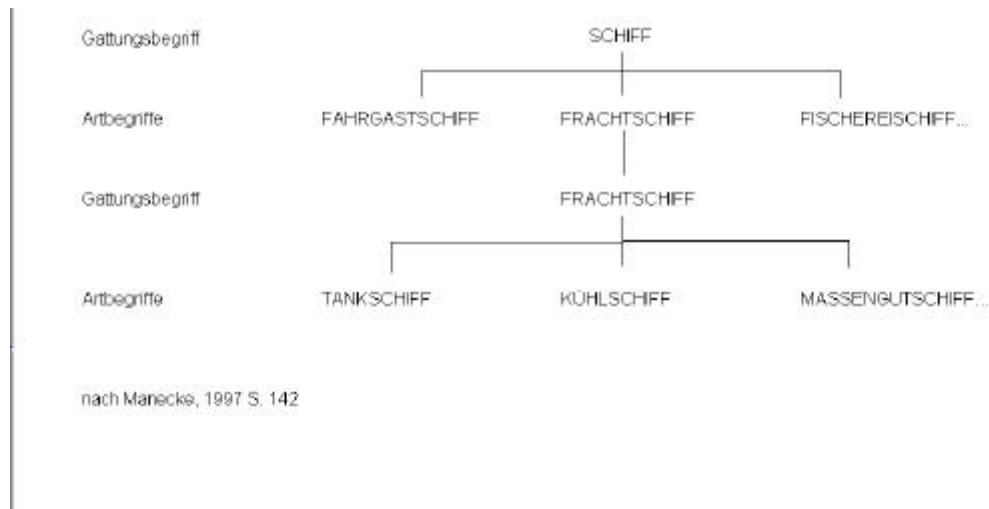


Abbildung 1 Beispiel einer analytischen Klassifikation

In obiger Abbildung sind den Hauptbegriffen (Gattungsbegriffen) verschiedene Unterbegriffe (Artbegriffe) monohierarchisch zugeordnet. Jeder Unterbegriff hat nur einen Oberbegriff.

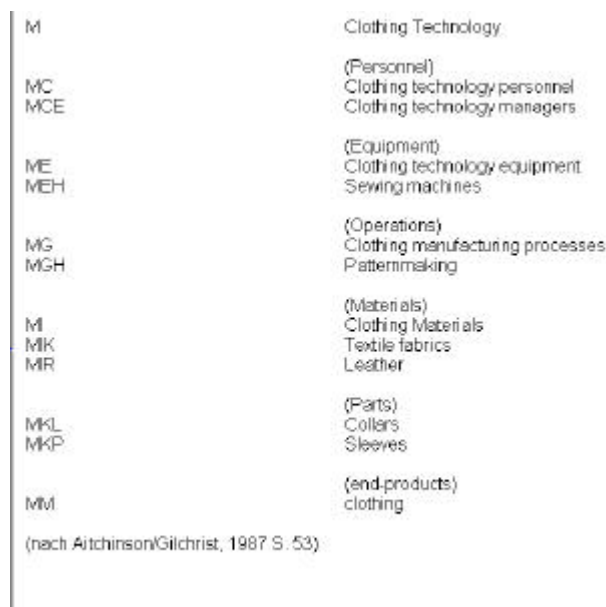


Abbildung 2 Beispiel einer facettierten Klassifikation

In der Abbildung werden einzelne Gruppen der Hauptgruppe „M“ „Clothing Technology“ durch bis zu drei Buchstaben dargestellt. Je mehr Buchstaben, desto tiefer geht die

Begrifflichkeit. Einzelne Gruppen werden in Klammern über die feinere Gliederung geschrieben.

Allerdings werden heutzutage, bedingt durch die verbesserte Rechnerleistung, beide Formen zur Erschließung von Wissensquellen benützt, da es auf deren Zweckmäßigkeit, nicht mehr auf die Technik ankommt (Manecke, 1997 S. 147).

Thesauri arbeiten mit natürlichsprachlichen Begriffen, während andere Dokumentations-sprachen mit künstlichsprachlichen aufgebaut sind. So führt ein Zugangsterminus in der Regel auf Nummern, Code oder eine Notation, während beim Thesaurus er wieder auf einen natürlichsprachlichen Ausdruck führt.

Künstliche Begriffe gibt es bei Thesauri meist nur in der Speichersprache.

Natürliche Sprachen sind mehrdeutig, das heißt, Aussagen sind auf verschiedene Weise interpretierbar. Künstliche Sprachen sind hingegen eindeutig. Ein eindeutiges Ausdrücken in natürlichen Sprachen ist möglich, aber meist mit größerem Aufwand verbunden (Soergel, 1971 S. 265).

Dies bemerke ich auch bei meiner praktischen Arbeit am Thesaurus. „Konversion“ kommt zum Beispiel in verschiedenen Fachgebieten (wie zum Beispiel Religion, Logik, Finanzwesen, Recht und Kernphysik) vor. Von all diesen Möglichkeiten kommt aber für unser Fachgebiet nur letzteres in Frage, da sie etwas mit Umweltvorgängen, nämlich Kernverfall, zu tun hat. Der Begriff ist nun genauer bestimmt.

Gleiches gilt für den Begriff „Einführung“ der zunächst sehr Allgemein gehalten ist. Man kann ihn aber noch weiter spezifizieren, indem er in „Einführung von Pflanzenarten“ und „Einführung von Tierarten“ aufgeteilt wird. Ebenso wird er Begriff „Eingriff“ in „Eingriff in die Raumordnung“ und „Eingriff in Natur und Landschaft“ untergliedert.

Soergel (1971 S. 245) weist darauf hin, dass zunächst die bestehenden Dokumentations-sprachen und Klassifikationssysteme zu koordinieren und konvertibel zu machen sind, ehe an die Schaffung einer universellen Dokumentations-sprache gedacht werden kann.

Klassifikationen benutzen hierarchische Beziehungen zur Systembildung, während beim Thesaurus sämtliche Beziehungsarten gleichberechtigt vorkommen können.

Klassifikationen müssen viele verschieden Klasseneinteilungen vornehmen, um so allgemein wie möglich zu sein, während Thesauri nur die Beziehungen der begrifflichen Einheiten darstellen, welche für sie relevant sind.

Bei den hierarchischen Klassifikationen ist noch anzumerken, dass sie monohierarchisch und monodimensional geordnet sind, während facettierte Klassifikationen und Thesauri polyhierarchisch und polydimensional aufgebaut sind

Verglichen mit anderen Systemen wird der Thesaurus wesentlich stärker kontrolliert: Stich- und Schlagwortsysteme üben eine vollständige Vokabularzugangskontrolle und

Terminologiekontrolle aus. Jeder zur Indexierung benutzte Term muss innerhalb der Dokumentationssprache spezifiziert werden. Schließlich stellen Schlag- und Stichwort-systeme nach einem systematischen Verfahren Beziehungen zwischen einzelnen Begriffen dar.

Es gibt aber auch fließende Übergänge zwischen den Formen (Wersig 1978, S. 30-32).

Beim Thesaurus ohne Vorzugsbenennung gelten alle Begriffe als gleichwertig, das heißt, sie können zum Indexieren und Retrieval benützt werden. Die Äquivalenzklasse wird durch eine Nummer (Notation) verdeutlicht, die für den Benutzer nicht sichtbar ist, und ordnenden Charakter hat (Burkart, 1997 S. 168).

Beim Thesaurus mit Vorzugsbenennung wird ein Klassenelement der Äquivalenzklasse als bevorzugt eingestuft, auf andere wird verwiesen (Burkart, 1997 S. 169).

Um diese Art der Thesauri soll es nun gehen.

3.3 Verschiedene Thesaurusdefinitionen

Nach ISO 2788 ist ein Thesaurus

das Vokabular einer kontrollierten Indexierungssprache, formal organisiert so dass a priori Beziehungen zwischen Begriffen (zum Beispiel „breiterer“ und „engerer“ Begriff“) deutlich gemacht werden.

Ein kontrollierte Begriffssammlung aus einer natürlichen Sprache ausgewählt, die dazu benützt wird, in abstrakter Form Dokumentssubjekte zu veranschaulichen (ESI, 2001).

In ähnlicher Weise äußert sich DIN 1463, die von ISO abgeleitet wurde:

„Ein Thesaurus (...) ist eine geordnete Zusammenstellung von Begriffen, und ihren (...) Bezeichnungen, die in einem Dokumentationsgebiet zum Indexieren, Speichern und Wiederauffinden dient.“ (S. 22)

„In einem Thesaurus wird eine Reihe von Beziehungen zwischen Begriffen und ihren Bezeichnungen dargestellt. Dient der Vokabularkontrolle und der terminologischen Kontrolle (Spree, 2003).“

3.4 Verschiedene Relationstypen

Tabelle 1 DIN 1462 gibt für die Beziehungen zwischen den einzelnen Begriffen folgende Kürzel an:

BF - Benutzt für Synonym
BS - Benutze Synonym
OB - Oberbegriff
UB - Unterbegriff
VB - Verwandter Begriff

Tabelle 2 ISO 2788 nennt folgende Begriffe:

UF - Used for
USE/SYN Use synonym
BT - Broader Term)
NT - Narrower term
RT - Related term
TT - Top term

(Wikipedia, 2003).

Riekert (1996, S 2) erwähnt noch zwei weitere Begriffsbeziehungen: BK (Synonymkombinationen) und OOB für (Ober-Oberbegriff), dies entspricht dem englischen TT (Top Term).

Im folgenden werden die einzelnen Abkürzungen beispielhaft erklärt. Es werden zunächst die deutsche Form, in Klammern die englische Version wiedergegeben.

Im voraus muss noch bemerkt werden, dass die Beziehungen zwischen Begriffen

„immer reziprok wiedergegeben werden, d. h. nicht nur vom Nichtdeskriptor auf den Deskriptor verweisen, sondern zu jedem Deskriptor alle Nichtdeskriptoren aufweisen, um den ganzen Umfang der Äquivalenzklasse aufzuzeigen“ (Burkart, 1997 S. 171).

3.4.1 Äquivalenzrelation

BS bedeutet „Benutztes Synonym“ (engl. USE/SYN „use synonym“) und BF „Benutzt für“ (engl. UF „used for“). Das Synonym Sonnabend verweist auf den Deskriptor Samstag.

Sonnabend BS Samstag

Samstag BF Sonnabend.

Weitere Beispiele aus meiner praktischen Arbeit: so verweisen die Abkürzungen „ÖPNV“ auf „Öffentlicher Personennahverkehr“, „USA“ auf „United States of America“ „UVP“ auf „Umweltverträglichkeitsprüfung“, um nur einige zu nennen. So sieht dann die Schematische Schreibweise mit englischsprachigen Abkürzungen aus:

ÖPNV UF Öffentlicher Personennahverkehr

Öffentlicher Personennahverkehr USE ÖPNV

USA UF United States of America

United States of America USE USA

UVP UF Umweltverträglichkeitsprüfung

Umweltverträglichkeitsprüfung USE UVP

Ich habe mich für diese Form entschieden, weil „UVP“, „USA“ und „ÖPNV“ die gängigeren Begriffe sind, intern aber auch auf deren Langformen verwiesen werden soll.

3.4.2 Generische Relation

Eine generische Anordnung liegt dann vor, wenn ein Narrower Term alle Eigenschaften des Broader Term hat und mindestens eine zusätzliche Eigenschaft. Ein Beispiel aus der Praxis:

„Vogel“ NT „Singvogel“ NT „Rotkehlchen“

„Rotkehlchen“ BT „Singvogel“ BT „Vogel“

3.4.3 Partitive Relation

Die partitive Relation wird auch Bestandsrelation genannt. Ein übergeordneter Begriff entspricht einem Ganzen, der untergeordnete Begriff ist ein Teil des Ganzen.

OB bedeutet Oberbegriff (engl. BT Broader Term) und UB Unterbegriff (engl. NT Narrower Term).

Baum UB Baumstamm

Baumstamm OB Baum.

3.4.4 Assoziationsrelation

VB bedeutet Verwandter Begriff (RT Related Term).

Obst VB Obstbaum

Obstbaum VB Obst

Die Assoziationsrelation ist weder hierarchisch noch äquivalent. Sie wird in beide Richtungen verwendet und es kommt vor, dass in ihr alle in irgendeiner Form verwandten Wörter versammelt werden, was zu unübersichtlichen Beziehungen führen kann, oder es werden alle auf einer Beziehungsebene befindlichen Begriffe miteinander über VB verlinkt, dies kann zur Redundanz führen, da diese Bezüge auch schon in der Hierarchischen Beziehung repräsentiert sein können.

3.4.5 Begriffskombination

BK steht für „Benutzte Kombination“ und KB „Kombinationsbegriff“. Sie entstehen dann, wenn man aus zwei Deskriptoren einen Begriff zusammensetzt. Dieser gehört dann zu den Non - Deskriptoren. Da aber seine Äquivalenzklasse eine Schnittmenge der beiden Deskriptoren darstellt, die es im Thesaurus eigentlich nicht gibt, wurde dieser Begriff eingeführt.

Geldbriefträger BK Briefzusteller + Geld

Briefzusteller KB Geldbriefträger

Geldzustellung KB Geldbriefträger.

Wersig (1985, S. 118) zerlegt den reziproken Eintrag in drei Teile.

Geldbriefträger BK Briefzusteller + Geldzustellung

Briefzusteller BIK (Benutzt in Kombination) Geldzustellung BFK (Benutzt für Kombination) Geldbriefträger.

Geldzustellung BIK Briefzustellung BFK Geldbriefträger.

Die Einträge sollten entweder alternierend dargestellt, oder mit Indikatoren versehen werden, um Verwirrungen zu vermeiden (Burkart, 1997 S. 170-175).

Diese zusammengesetzten Begriffe, werden auch Komposita genannt. Ihre Handhabung in Thesauri ist schwierig, da sie zerlegt werden müssen.

Erfreulicherweise ist mein Ausgangsmaterial so gestaltet, dass sie sehr selten vorkommen. Ansonsten kommen sie im Deutschen im Vergleich etwa zum Englischen, häufig vor. Beispielhaft sei hier die „Donaudampfschiffahrtsgesellschaft“ genannt.

Anstelle einer Notation geben einige Thesauri neben den Oberbegriffen Oberste Oberbegriffe (TT Top Term) an. Diese werden aber nicht hierarchisch verwendet, sondern haben nur erläuternden Charakter (Burkart, 1997 S. 176).

3.5 Probleme bei der Thesauruserstellung

Nachfolgend werden einige Probleme angesprochen, mit denen man sich bei der Erstellung eines Thesaurus mehr oder weniger konfrontiert sieht.

3.5.1 Synonymproblem

Synonyme sind unterschiedliche Wörter, die jedoch die gleiche Bedeutung haben. Wersig (1978, S 49-52) unterscheidet folgende Fälle.

vollständige Identität: Wörter werden anders geschrieben oder es gibt Kurzformen von ihnen.

Telephon - Telefon

Bus - Omnibus

Denotative Identität: beide Begriffe bezeichnen die gleichen Objekte, haben aber unterschiedliche Bedeutungsnuancen.

Datenverarbeitungsanlage - Computer

Presse-, und Informationsamt der Bundesregierung - Bundespresseamt

Lebensmittel - Nahrungsmittel

Äquivalente Verwendung: bei geringfügigen bedeutungsmäßigen oder denotativen Unterschieden.

Holland - Niederlande

Abessinien - Äthiopien

Krebs (umgangssprachlich) - Karzinom (fachsprachlich)

Weiterhin nennt er noch bei den Quasi Synonymen:

Begriffliche Überschneidungen

Dauer - Zeit

Gleichsetzung von verschiedenen Erscheinungsformen

Rauheit - Glätte

kontradiktorischen Gegensätzen

Metall - Nichtmetall

von Ober-, und Unterbegriffen, wenn sie nur durch eine Stufe getrennt, und inhaltlich gleich sind.

Nachschlagewerk - Lexikon

Schließlich die Zuordnung spezielle Unterbegriffe zu einem Oberbegriff, wenn die Hierarchieebene für den Thesaurus nicht relevant ist:

Pferd - Gaul (ist Synonym)

aber Pferd – Rennpferd (getrennt, z. B. im Sportthesaurus).

Zur Unterstützung können diese Begriffe semantisch zerlegt werden, aber sie müssen durch logische Kombination wieder zum Ausgangsbegriff zusammengefügt werden können (Wersig, 1978 S. 58). So wird zum Beispiel der Begriff „Eisenbahn“ nicht in „Eisen“ und „Bahn“ sondern in „Schienenbahn“ und „Überlandverkehr“ zerlegt (Wersig, 1978 S. 59).

3.5.2 Polysemproblem

Homonyme (gleich klingende oder geschriebene Wörter, die unterschiedliche Bedeutung haben) werden von Polysemen unterschieden. Es gibt Homophone, deren Ähnlichkeit ist nur lautlich bedingt und Homographen, bei denen die Ähnlichkeit schriftlich bedingt ist.

Bei echten Homographen lässt sich die gleiche Schreibweise historisch nachvollziehen, während bei unechten Homographen dies nicht der Fall ist.

Erfuhr ein Wort durch Übernahme von einem Sprachbereich in einen anderen einen Bedeutungswandel nennt man es Polysem. Diese werden entweder als nicht Thesaurus-relevant betrachtet oder durch Zusätze näher erklärt (Wersig, 1978 S 67).

3.5.3 Deskriptoren

Deskriptoren in einem Thesaurus sind Begriffe, die bevorzugt verwendet werden. Non-Deskriptoren führen zu Deskriptoren und sind nicht bevorzugt.

Wersig (1978, S 86) definiert Deskriptoren so:

“Als Deskriptor soll jede Benennung einer natürlichen Sprache (oder eine im Rahmen eines Thesaurus analog gebildete Benennung) gelten, die für Indexierung und Retrieval uneingeschränkt als Repräsentation eines thesaurusmäßig gebildeten Begriffssatzes verwendet werden kann.“

3.5.4 Non - Deskriptoren

Non - Deskriptoren können nicht als Thesaurusvokabular verwendet werden,

„führen aber durch vom Thesaurus vorgegebene Strukturen auf eine verbindliche Repräsentation mindestens eines Begriffssatzes“ (Wersig 1978, S. 102).

3.6. Arten von Thesaurussoftware

Nohr (1993) unterscheidet zwei verschiedenen Arten von Thesaurussoftware:

3.6.1. Eigenständige Thesaurussoftware

Sie ermöglicht das Verwalten und Erstellen von terminologischen Anwendungssystemen und ist zunächst nur für diese spezielle Aufgabe bestimmt. Da Thesauri aber nur in Umgebungsanwendungen, wie Retrievalprogrammen, ihren vollen Nutzen entfalten können, ist eine Anbindung an dieselben wichtig. Von Vorteil ist aber die genaue Fokussierung auf Entwicklung und Aufbau des Thesaurus.

3.6.2. Integrierte Thesaurussoftware

Sie ist nur innerhalb ihrer eigenen Umgebung anwendbar und besitzt meist auch keine Konvertierungsprogramme zu anderen Umgebungen. Da integrierte Thesaurussoftware viele verschiedene Komponenten umfasst, wie Erfassung, Thesaurus, Retrieval und Druck, dieselben alle aufeinander abgestimmt werden müssen, kann sie nicht so ausgefeilte Instrumente zur Thesaurusentwicklung wie eigenständige Thesaurussoftware anbieten. Auch wirkt sich jede Änderung einer Komponente im System auf eine andere aus, so dass integrierte Software immer das Ganze, nicht eine einzelne Komponente, im Blick haben muss.

3.6.3. Anforderungen an eine Thesaurussoftware

Nach Burkart (1997, S. 179) sollte Thesaurussoftware folgende Anforderungen erfüllen: automatische Erzeugung reziproker Einträge, Verhinderung von Doubletten, Plausibilitätskontrollen um Ringverweisungen zu unterbinden. Alle Änderungen sollten automatisch durchgeführt werden.

Bei „Thew 32“ handelt sich es nach den vorangegangenen Ausführungen um eine eigenständige Thesaurussoftware. Ich habe mich dafür entschieden, da die Thesauruserstellung zunächst im Vordergrund stehen soll. Später können dann immer noch andere Komponenten wie zum Beispiel eine Suche hinzugefügt werden. Dies ist aber nicht unbedingt nötig, da in „Thew 32“ sowohl über eine Suchfunktion, als auch durch klicken im systematischen Thesaurus, gesucht werden kann.

„Thew 32“ erfüllt auch diese Anforderungen. So werden doppelte Einträge gelöscht, Ringverweise nicht zugelassen und die Einträge wechselseitig vorgenommen.

4. Verschiedene Umweltthesauri

Nachfolgend werden einige wichtige Umweltthesauri näher beschrieben.

4.1 Umweltthesaurus des Umweltbundesamtes

Batschi (1994 S. 57-62) gibt einen Überblick über die Geschichte des Umweltthesaurus des Umweltbundesamtes.

4.1.2 Aufgaben des Umweltbundesamtes Berlin

Das Umweltbundesamt Berlin wurde 1974 mit dem Ziel gegründet, das Bundesministerium für Umwelt, Naturschutz und Reaktorsicherheit wissenschaftlich zu unterstützen, ein Informationssystem zur Umweltplanung und eine zentrale Umweltdokumentation zu entwickeln und zu unterhalten, die Untersuchungen des Ministeriums vorzubereiten und zu unterstützen, sowie die Umweltuntersuchungen der Behörden zu koordinieren und schließlich, die Öffentlichkeit zu informieren.

4.1.3 Umweltdatenbanken des Umweltbundesamtes

Aus diesen Gründen wurden mehrere Umweltdatenbanken entworfen und entwickelt, von denen sämtliche (MONUFAKT ausgenommen) seit 1984 online verfügbar sind:

Tabelle 3 Umweltdatenbanken des Umweltbundesamtes

ULITAD/ULIT	Umweltliteraturdatenbank, bietet > 200 000 deutschsprachige Titel an.
UFORDAT/UFOR	Umweltforschungs- und Entwicklungsdatenbank für Projekte, bietet > 32 000 Projekte und 8 8000 Institutionen.
URDB	Umweltrechtsdatenbank: 4 500 Gerichtsurteile, 11 000 Gesetze (auf Bundes-, und Länderebene), 350 internationale Verträge, 2 5000 Gesetze der EU.
MONUFAKT	Datenbank über Gebäudeschäden, die durch Umweltprobleme verursacht wurden, bietet > 3 100 Titel.

Diese Datenbanken wurden entwickelt, da nur ein kleiner Teil deutschsprachige Literatur in internationalen Datenbanken vorhanden war und die Dokumentationsstellen verschiedene Indexierungsregeln und –werkzeuge benutzten.

4.1.4 Der Umweltthesaurus des Umweltbundesamtes

Es war auch kein Thesaurus zur Erschließung der Dokumente vorhanden. Zunächst wurde versucht, die verschiedenen Begriffe der Dokumentationszentren in einem Thesaurus zusammenzuführen. Dies misslang aber, da zu viele Indexierungsterme erfunden wurden, um eine Hierarchie aufzubauen, so dass der Thesaurus unhandlich wurde. Dann wurden als kontrollierte Begriffe Indexierungsterme aus den Datenbanken und freie Begriffe aus wichtigen Dokumenten verwendet. Nach einer Zugangskontrolle wurden sie als kontrollierte Begriffe in den Thesaurus aufgenommen und in eine hierarchische Abfolge gebracht. Es wurden nur wirklich gebräuchliche Begriffe in Betracht gezogen.

Es wurde eine Umweltklassifikation entwickelt und ein Umweltthesaurus. Letzterer zeichnet sich durch einen polyhierarchischen, multilingualen Aufbau und der Benutzung von OB, UB, Synonyme und Quasisynonyme und BK aus. Er besteht aus 8 100 Begriffen und 22000 Non - Deskriptoren. Der Geographische Thesaurus, der in einem separaten Dokument gehalten wird, besteht aus mehr als 3600 Deskriptoren und 2800 Non-Deskriptoren. Momentan ist er noch zweisprachig (Deutsch-Englisch) eine weitere Mehrsprachigkeit, insbesondere in Osteuropäischen Sprachen, wird aber angestrebt.

Als Software wird aDIS (adaptable Documentation and Information System) verwendet, die Vergleiche, Rechtschreibprüfungen, Kompositazerlegung, Synoymerkennung und deren Austausch durch Deskriptoren, Termgewichtung durch Feststellung der Häufigkeit des Vorkommens im Text, die Hierarchieebene des Terms sowie einen richtigen Klassifizierungscode anwenden kann. Mit ihr können 80 % der Terme automatisiert erkannt werden.

4.1.5 Aufgaben des Umweltbundesamtes auf nationaler und internationaler Ebene

Ferner werden auf nationaler Ebene Datenobjekte verschiedener Bundesländer und Bücher deutscher Umweltbehörden indiziert, Verzeichnisse von Bibliographien, jährliche Berichte, Kataloge und Veröffentlichungen vorbereitet.

Auf internationaler Ebene wird in einem Projekt der österreichische Thesaurus für Datenquellen über Umweltdaten zu einem multilingualen Instrument (Deutsch, Englisch, Französisch und Italienisch) ausgebaut. Die Ergebnisse sollen dem Europäischen Umweltbundesamt zu Gute kommen, um einen mehrsprachigen Umweltthesaurus aller Sprachen der Europäischen Union zu entwickeln. Dieser wird im Folgenden vorgestellt.

4.2 General Multilingual Environmental Thesaurus (GEMET)

GEMET ist das Indexierungs-, Retrieval und Kontrollwerkzeug des ETC/CDS und des EEA. Es wurde angestrebt, einen allgemein gültigen Thesaurus in natürlicher Sprache

zu schaffen, spezielle Thesauri und Deskriptorensysteme wurden nur strukturell und auf oberer Terminologieebene benutzt.

Folgende multilinguale Dokumente wurden zusammengeführt:

Eine Auswahl des Umweltthesaurus des Umeltbundesamtes, vollständig der „Thesaurus Italiano per l’Ambiente (TIA) des Consiglio Nazionale delle Ricerche (CNR), Multilingual Environment Thesaurus (MET) des Nederlands Bureau voor Onderzoek Informatie (NBOI), EnVoc Thesaurus des UNEP Infoterra, Thesaurus de Medio Ambiente des Ministerio de Obras Publicas, Transportes y Medio Ambiente (MOPTMA), Lexique environment –Planète des Ministère de l’environnement, schließlich Deskriptoren wichtiger Dokumente des EEA und die Terminologie des ECT und EIONET (Netzwerk des ECT), sowie Deskriptoren des „Thesaurus Eurovoc“ des Europäischen Parlamentes.

Diese 6562 Begriffe wurden in einen Klassifikationsrahmen in 3 Hauptgruppen, 30 Gruppen und 5 instrumentelle Gruppen angeordnet. Jeder Deskriptor wurde hierarchisch angeordnet und besitzt einen Top Term. Um einen thematischen Zugang zum Thesaurus zu ermöglichen, wurden 40 Themen mit dem EEA vereinbart. Jeder Deskriptor wurde so vielen Themen zugeordnet wie notwendig. Nun kann der Benutzer den Thesaurus durch hierarchische, thematische, oder alphabetische Listen betreten. Da es nicht möglich war, hierarchische „vertikale“ Beziehungen aufzubauen, wurden starke „horizontale“ (RT, Related Terms) Beziehungen zwischen den Begriffen aufgebaut. Jedem Deskriptor der systematischen Liste wurde eine fortlaufende Nummer gegeben, um es dem Benutzer zu ermöglichen, den Deskriptor in verschiedenen Listen zu erkennen. Diese Nummer ist nur ein neutraler Indikator des Platzes, auf dem der Begriff gerade steht.

GEMET folgt der ISO Norm für mono-, und multilinguale Thesauri. Entwickler des GEMET können auf das Software Programm THESmain des Technischen Büros Hermann Stallbaumer, Wien zurückgreifen, während Benutzer THESshow benutzen können. GEMET ist aber auch auf CD-ROM und als PDF- Dokument erhältlich, welche die bisherigen Druckversionen ersetzen.

Die Software THESshow dient zum anzeigen der mit Hilfe von THESmain erstellten Datenbeständen. Diese sind in den Sprachen Englisch bzw Deutsch verfügbar und können hierarchisch bzw alphabetisch angezeigt werden. Terme können gesucht und in einen Datenkorb abgelegt werden. (Legat, 1999).

Im GEMET können Deskriptoren entweder generisch angeordnet (ein Narrower Term hat alle Eigenschaften des Broader Term und mindestens eine zusätzliche Eigenschaft), oder Ganz-Teil Beziehungen sein, (ein Narrower Term muss ein Teil des Broader Terms sein). Besitzt ein Term beide Möglichkeiten ist er polyhierarchisch angeordnet und ist gleichwertig. Non - Deskriptoren (Synonyme von Deskriptoren) gehören zur gleichen Gruppe wie der Deskriptor und führen den Benutzer zum Deskriptor, der bevorzugten Form. Im GEMET wurde durchweg Singular verwendet, nur einige englische

Begriffe wurden, um die Bedeutung nicht zu verändern, im Plural belassen. Alternative Schreibweisen und Varianten wurden als Non - Deskriptoren berücksichtigt, aber nur im Englischen (GEMET 2001, I-V).

5. Eingeschlagener Realisierungsweg

Nun wird der Realisierungsweg vorgestellt.

Einleitend möchte ich bemerken, dass Studenten im Fach „Inhaltliche Erschließung“ bei Herrn Prof. Nohr im Studiengang Informationsmanagement der HdM im Sommersemester 2000 mit der Software „Thew 32“ Thesauri entworfen haben. Beispielhaft seien hier der Thesaurus des menschlichen Skeletts von Jasmin Wendt und Gundula Marscke, sowie der Fahrradthesaurus von Katja Richter und Matthias Kissler genannt. Beide können für das Aussehen des zu erstellenden Thesaurus wichtige Hinweise geben. (Nohr, 2000).



Abbildung 3 Beispiele für Thesauri, welche mit „Thew 32“ realisiert wurden

5.1 Import der Schlagworte und verschiedene Ordnungsarbeiten

Aus der vom Autor erstellten Bibliotheksdatenbank (Müller, 2003) wurden die Schlagworte in ein Textdokument importiert, um sie für den Vergleich mit GEMET vorzubereiten. Es wird zunächst ein Schlagwortvergleich zwischen den beiden gemacht. Dann werden sie nach der Ordnung des GEMET geordnet und später mit dem „thew 32“ Programm realisiert.

```
3;"Abfall"  
4;"Abfallbehandlung"  
5;"Abfallbeseitigung"  
6;"Abfallvermeidung"  
7;"Abfallwirtschaft"  
8;"Abfluß"  
9;"Abgas"  
10;"Abgrabung"  
11;"Abtragung"  
12;"Abwasser"  
13;"Abwasserbehandlung"  
14;"Adressen"  
15;"Agrar"  
16;"Agrar-"  
17;"Akustik"  
18;"Allgemeines"  
19;"Alternative"  
20;"Alternativen"  
21;"Altlast"  
22;"Amphib"  
23;"Anropologie"  
24;"Archäologie"  
25;"Architektur"  
26;"Arten"  
27;"Artenschutz"  
28;"Atmosphäre"  
29;"Ausland"  
30;"Ausleitung"  
31;"Auswirkung"  
32;"Auto"  
33;"Autobahn"
```

Abbildung 4 Importierte Schlagwortliste aus Bibliotheksdatenbank

Zunächst müssen noch Duplikate ausgesondert werden.

Sollen im Thesaurus sowohl Singular als auch Pluralformen vorkommen? Im GEMET wurde durchweg Singular verwendet, deshalb möchte ich es auch so halten. Entsprechend äußern sich auch Aitchison und Gilchrist (1987, S. 14). Franzosen und Deutsche tendieren zum Singular, außer wenn Singular und Plural verschiedene Bedeutungen haben. Während im Englischen sowohl Singular als auch Plural nach Regeln und Standards, die in diesem Gebiet gelten, verwendet werden.

Kommen zwei Begriffe, ein wissenschaftlicher und ein gemeingebräuchlicher, im Thesaurus vor, soll der in der Organisation meist verwendete benutzt werden. Auf die wissenschaftliche Form soll verwiesen werden (Aitchison, Gilchrist, 1987 S.18). Zum Beispiel wird der Begriff „ÖPNV“ zum Deskriptor, auf „Öffentlicher Personennahverkehr“ wird entsprechend als Non - Deskriptor verwiesen.

```
!86;"Psychologie"  
!87;"Pumpwerk"  
!88;"Radioaktivität"  
  
!90;"Raumordnung"  
!91;"Raumordnungsverfahren"  
!92;"Raumplanung"  
!93;"Recht"  
!94;"Recycling"  
!95;"Region"  
!96;"Regionalplanung"  
!97;"Renaturierung"  
!98;"Reptil"  
!99;"Resource"  
!00;"Resources"  
!01;"Ressource"  
!02;"Rheinland-Pfalz"  
!03;"Richtlinie"  
!04;"Risiko"  
!05;"Rohstoff"  
!06;"Rohstoffe"  
!07;"Saarland"  
!08;"Sachsen"  
!09;"Sachsen-Anhalt"  
!10;"Sammlung"  
!11;"Säugetier"  
!12;"Schadstoffe"  
!13;"Schadstoffe"  
!14;"Schadstoffvermeidung"  
!15;"Schiene"  
!16;"Schlagworte"  
!17;"Schleswig-Holstein"  
!18;"Schutz"  
!19;"Schutzgebiet"  
!20;"Schutzverordnung"  
!21;"Schweiz"  
!22;"Scoping"  
!23;"See"  
!24;"Sicherheit"  
!25;"Siedlung"  
!26;"Siedlungsentwicklung"  
!27;"Siedlungsökologie"  
!28;"Siedlungsstruktur"  
!29;"Skisport"  
!30;"Solar"
```

Abbildung 5 Entfernen von Duplikaten und Pluralformen

Ferner wird anhand einer aktuellen, gedruckten Schlagwortliste eine Überprüfung und gegebenenfalls eine Angleichung vorgenommen. Ebenso werden für Abkürzungen die ausgeschriebenen Worte zum Beispiel Umweltverträglichkeitsprüfung für UVP eingefügt. Mit der Such- und Ersetzfunktion von Word wurden die Sonderzeichen wie ; , „ und Nummern entfernt. Die Liste sieht dann folgendermaßen aus:

Abfall
 Abfallbehandlung
 Abfallbeseitigung
 Abfallvermeidung
 Abfallwirtschaft
 Abfluss
 Abgas
 Abgrabung
 Abtragung
 Abwasser
 Abwasserbehandlung
 Adresse
 Agrar-
 Akustik
 Allgemeines
 Alternative
 Altlast
 Amphib
 Anthropologie
 Archäologie
 Architektur
 Arten
 Artenschutz
 Atmosphäre
 Atomenergierückhaltung
 Ausland
 Ausleitung
 Auswirkung
 Auto
 Autobahn
 Baden-Württemberg
 Bauen
 Baugesetz
 Bauland
 Bauleitplan
 Bauleitplanung
 Baum
 Baustoff
 Baustruktur
 Bayern
 Bebauungsplan
 Begleitplanung
 Begrünung
 Beispiel
 Belastung

Abbildung 6 Entgültige Schlagwortliste

Auf gleiche Weise wird mit den GEMET Schlagworten verfahren: sie werden in ein Worddokument kopiert, alle unnötigen Sonderzeichen entfernt, ebenso alle Duplikate. Sodann wird diese Liste in eine Access Datenbank kopiert.

(Kunst) handwerksunternehmen: §
 Abbau: §
 Abbau der Ozonschicht: §
 Abbau von natürlichen Ressourcen: §
 Abbaubarkeit: §
 Abbauprodukt: §
 Abbruchmüll: §
 Abbruchunternehmen: §
 Abdeckung: §
 Abdichtung: §
 Abfackelung: §
 Abfall: §
 Abfallabgabe: §
 Abfallablagerung: §
 Abfallart: §
 Abfallassimilationsfähigkeit: §
 Abfallaufkommen: §
 Abfallbehälter: §
 Abfallbehandlung: §
 Abfallbehandlungsabwasser: §
 Abfallbehandlungsanlage: §
 Abfallbeschaffenheit: §
 Abfallbeseitigung: §
 Abfallbeseitigung in den Boden: §
 Abfallbeseitigungsgesetz: §
 Abfallbeseitigungsunternehmen: §
 Abfallbilanz: §
 Abfallbörse: §
 Abfälle aus dem Gesundheitswesen: §
 Abfalleinführungsgenehmigung: §

Abbildung 7 Korrigierte GEMET Schlagwortliste

5.2. Vergleich zwischen GEMET und Bibliotheksschlagworten

Beide Schlagwortlisten werden in verschiedene Tabellen importiert, miteinander verknüpft, um in einer JOIN Abfrage doppelt vorkommende Wörter herauszufiltern. Eine Abfrage sortiert aus zwei Mengen nach bestimmten Kriterien eine Untermenge (hier die übereinstimmenden Begriffe) heraus und stellt sie in einer neuen Tabelle dar.

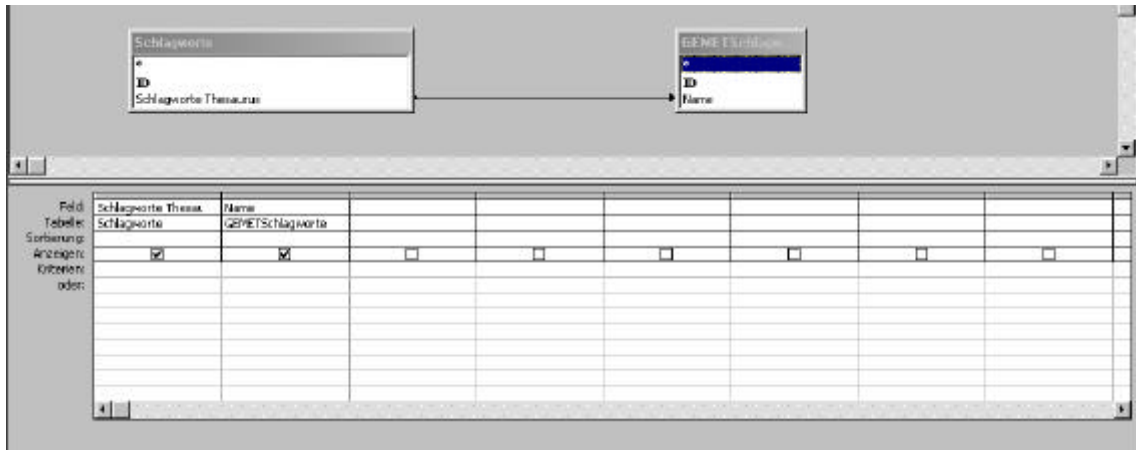


Abbildung 8 JOIN Abfrage mit dem Assistenten realisiert

Das Ergebnis sieht dann folgendermaßen aus:

Schlagworte Thesaurus	Name
Großstadt	Großstadt
Grundwasser	Grundwasser
Grundwasserschutz	Grundwasserschutz
Grünfläche	Grünfläche
Grünordnungsplan	
Grünordnungsplanung	
Grünplanung	
Gutachten	Gutachten
Güterindustrie	
Hafen	Hafen
Hamburg	
Haus	
Haushalt	Haushalt
Heimatkunde	
Hessen	
Hessen-Modell	
Hochwasser	Hochwasser
Hydraulik	Hydraulik
Hydrogeologie	Hydrogeologie
Hydrologie	Hydrologie
Immission	Immission
Indikator	Indikator
Individualverkehr	Individualverkehr
Industrie	Industrie
Information	Information
Infrastruktur	Infrastruktur
Infrastruktur	Infrastruktur
Innenstadt	Innenstadt
Insekt	Insekt
Kanal	
Karte	Karte
Kartierung	Kartierung
Kernkraft	
Kernanlage	

Abbildung 9 Vergleich von GEMET (links) und Bibliotheksschlagworten (rechts)

Wie sich zeigt, stimmen eine weitaus größere Anzahl der Schlagworte aus der Bibliotheksdatenbank mit den GEMET Schlagworten überein.

Tabelle 4 Eine Zählung der Schlagworthäufigkeiten ergibt folgendes Resultat:

	Anzahl
Bibliotheksschlagworte	400
Übereinstimmung	211

Daraus ergibt sich, dass von 400 Bibliotheksschlagworten mehr als die Hälfte mit den GEMET Schlagworten übereinstimmen.

5.3 Importieren der Schlagworte in Thesaurussoftware „Thew32“ und gegenseitige Verlinkung

Einführend möchte ich bemerken, dass der Import der Beziehungen von GEMET nach „Thew 32“ manuell erfolgte, da alle Versionen von GEMET als pdf – Dateien vorlagen und die Beziehungen nicht nach „Thew 32“ importiert werden können.

Zunächst müssen die Hauptgruppen in den Thesaurus aufgenommen werden, dies geschieht durch Import der Schlagworte nach „Thew 32“. Sodann werden sie durch Verlinkung mit der RT Beziehung miteinander verbunden. Hier werden die Begriffe „Forschung“, „Wissen“ und „Informationsgewinn“ mit dem Begriff „Wissenschaft“ verlinkt.

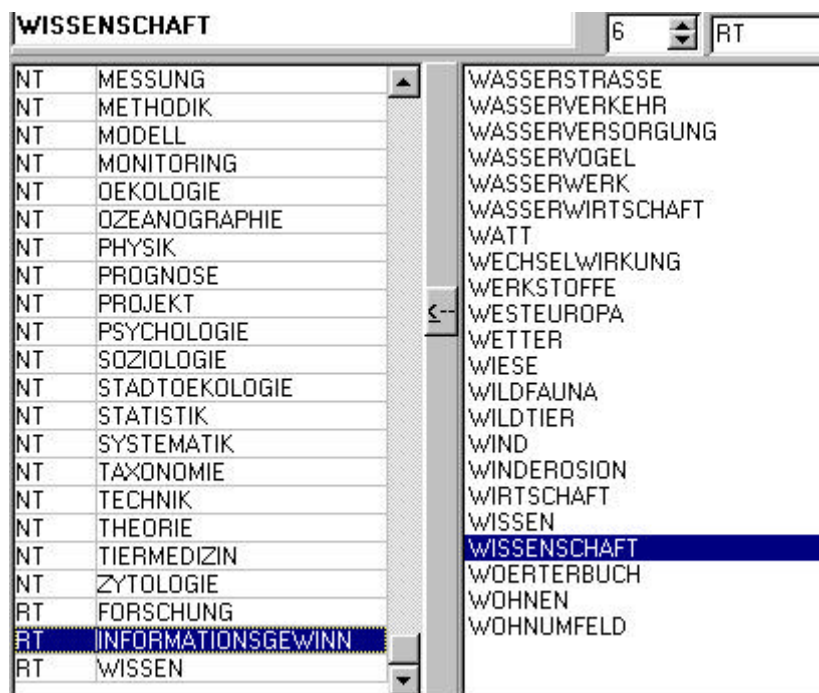


Abbildung 10 Verlinkung der Hauptgruppen des Thesaurus

Dann werden die häufig verwendeten Abkürzungen mit ihrer Hauptgruppe verlinkt und auf die Langform mit „UF“ verwiesen. Dies sieht dann wie in Abbildung 10 aus. Hier werden „Öffentlicher Personennahverkehr“ mit ÖPNV verlinkt.

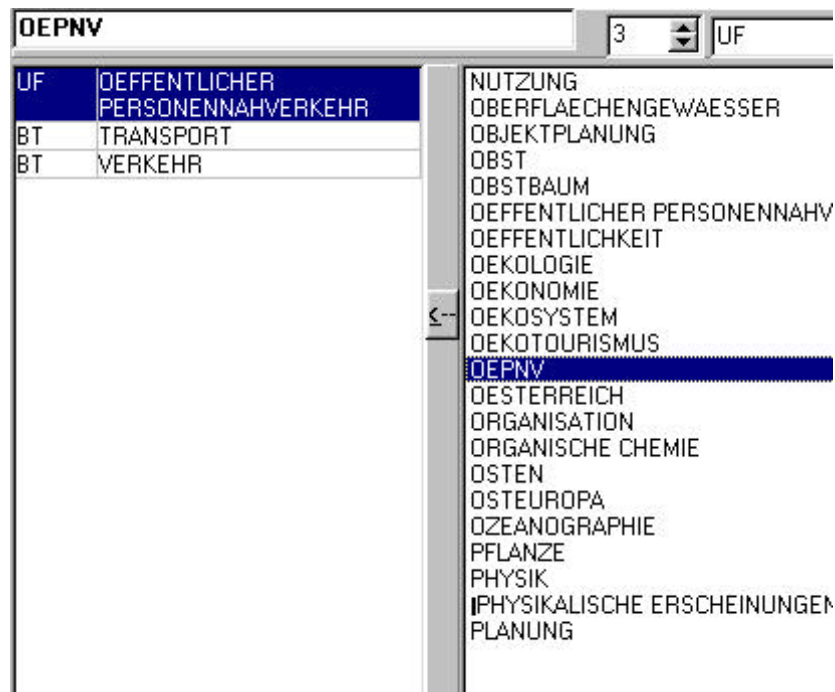


Abbildung 11 Verlinkung von Abkürzungen mit ihren Langformen

Im nächsten Schritt werden die einzelnen Hauptbegriffe als „Broader Term“ in den „Thew 32“ aufgenommen und die jeweiligen Unterbegriffe mit ihren Oberbegriffen verlinkt und umgekehrt. In Abbildung 12 wird der Begriff „Karte“ mit den Oberbegriffen „Ausbildung“, „Information“ und „Kultur“ verbunden. Der reziproke Eintrag bedingt gegenseitige logische Verlinkung, d.h. ist ein Begriff NT eines anderen, ist dieser automatisch BT des ersteren.

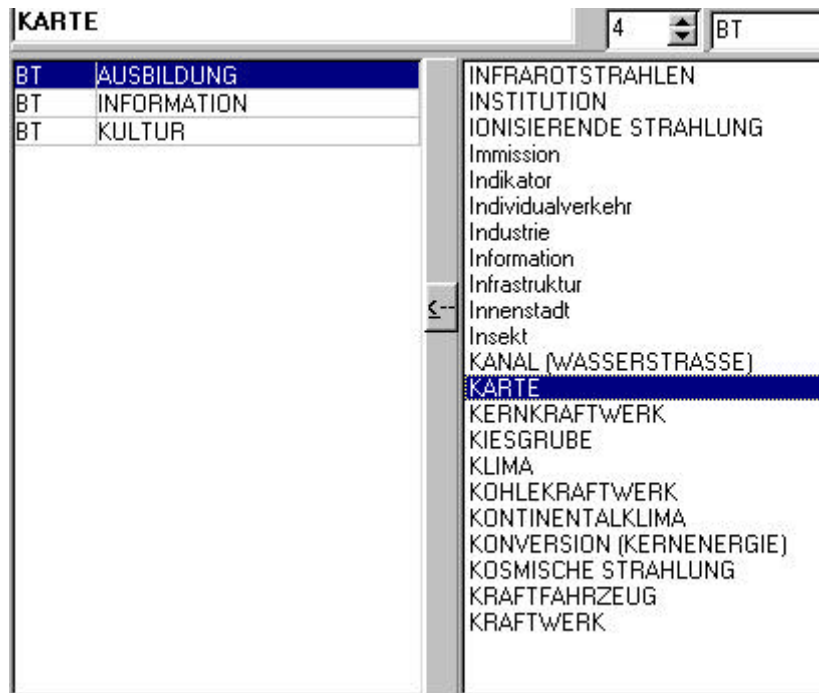


Abbildung 12 Verlinkung der Hauptbegriffe als „Broader Terms“

Sodann werden die übereinstimmenden Schlagworte aus ACCESS exportiert und in „Thew 32“ importiert.

Um eine Übereinstimmung mit dem GEMET zu gewährleisten, werden die Begriffe durch die Suchfunktion von Word gesucht, dann erfolgt die Verlinkung mit den Hauptschlagwörtern.



Abbildung 13 Schlagwortsuche im GEMET

Nachdem alle Begriffe mit den Broader Terms verlinkt sind, so wie es durch GEMET vorgegeben wird, werden die Schlagwörter, welche nicht im GEMET vorkommen ausgewählt und in „Thew 32“ importiert.

Schlagworte Thesaurus	Name
Genechnik	Genechnik
Geographie	Geographie
Geologie	Geologie
Geschichte	Geschichte
Geschwindigkeit	Geschwindigkeit
Geschwindigkeitsbog	
Gesellschaft	Gesellschaft
Gesetz	Gesetz
Gestaltung	Gestaltung
Gestein	Gestein
Gesundheit	Gesundheit
Gewässer	Gewässer
Gewässergüte	
Gewässerschutz	Gewässerschutz
Gewerbe	Gewerbe
Gewerbegebiet	
Golf	
Großbritannien	
Großstadt	Großstadt
Grundwasser	Grundwasser
Grundwasserschutz	Grundwasserschutz
Grünfläche	Grünfläche
Grünordnungsplan	
Grünordnungsplanung	
Grünplanung	
Gutachten	Gutachten
Güterindustrie	
Hafen	Hafen
Hamburg	
Haus	
Haushalt	Haushalt
Heimatkunde	
Hessen	
Hessen-Modell	

Abbildung 14 Auswählen der Schlagwörter, die nicht im GEMET vorkommen



Baden-Württemberg
Bauen
Baugesetz
Bauleitplan
Baustruktur
Bayern
Begleitplanung
Begründung
Beispiel
Belastung
Berlin
Bewertung
Biotiklima
Biotybernetik
Biotop
Biotopschutz
Brache
Brandenburg
Bremen
Brückenbau
Bund
Bundesland
Bundesland
Dammbauwerke
Daten
Definition
Denkmalschutz
Deutschland
Dorfenwicklung
Dorfneuerung

Abbildung 15 Fertige Liste der zu GEMET - nichtkonformen Schlagwörter

Später werden sie gleichfalls, diesmal auch aus dem Allgemeinwissen des Autors heraus, verlinkt. Schließlich werden Ergänzungen zu einigen Schlagworten mit Hilfe des GEMET angebracht. Dies kann aber nur bis zu einem gewissen Grade geschehen, da der Thesaurus möglichst einfach gehalten werden soll und es nicht unbedingt von Vorteil ist, wenn zu einem Begriff zehn Unterbegriffe aufgeführt sind, zu einem anderen hingegen nur einer. Dies geschieht folgendermaßen. Die erweiterten Begriffe werden mit ihrem Oberbegriff verbunden, dieser wird wieder gleichfalls mit einem Begriff auf der nächsthöheren Ebene verlinkt, letzterer dann mit den Hauptbegriffen. Beispielhaft sei hier Abbildung 16 angeführt.

RT Beziehungen werden nur dann verwendet, wenn Hauptbegriffe zusammen auf der höchsten Hierarchieebene sind, oder wenn sie sinnvoll erscheinen. So zum Beispiel ist „Sport“ mit „Freizeit“, „Erholung“, „Tourismus“ und „Fremdenverkehr“ verbunden, da sportliche Aktivitäten vornehmlich in der Freizeit ausgeübt werden. Ansonsten werden BT und NT Beziehungen verwendet.

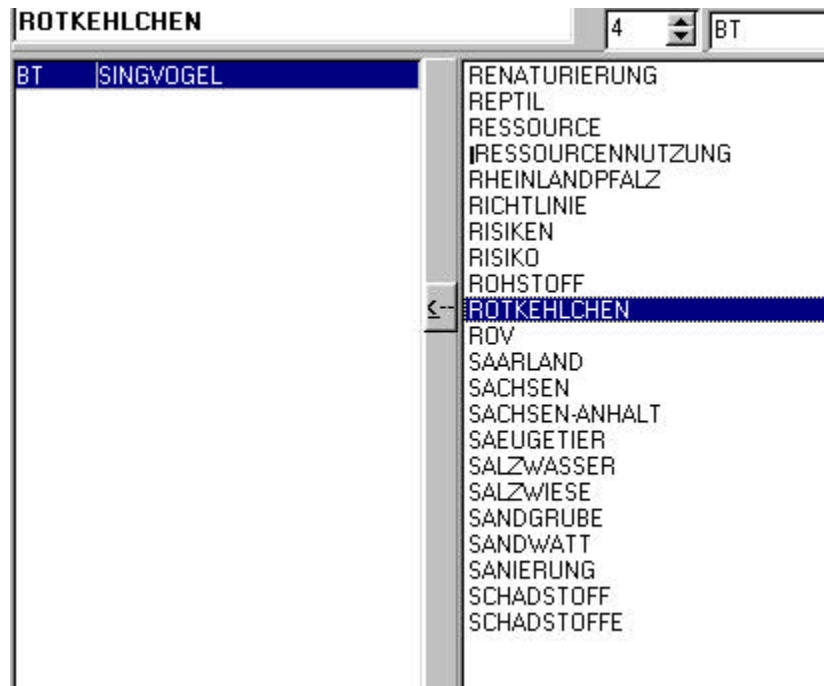


Abbildung 16 Ergänzen durch erweiterte Begriffe aus dem Thesaurus

In Abbildung 16 wird der erweiterte Begriff „Rotkehlchen“ mit seinem Oberbegriff „Singvogel“ verlinkt. „Singvogel“ ist dann mit „Vogel“ und „Vogel“ mit den Hauptbegriffen „Biologische Systeme und Prozesse“ und „Biosphäre“ verlinkt.

5.4 Abschließende Arbeiten am Thesaurus

Um den Thesaurus fertig zustellen, werden noch einige abschließende Arbeiten vorgenommen, die jetzt beschrieben werden.

5.4.1 Umwandlung importierter Begriffe in thesauruseigene

Da der Thesaurus importierte Begriffe zwar verlinkt, aber wenn man sie anklickt, keine weiteren Verlinkungen angezeigt werden, müssen sie durch den Befehl „Modify reference/note“ aus dem Editmenü zu thesauruseigenen Begriffen gemacht werden. Diese erkennt man an ihrer Großschreibung, da importierte Terme klein geschrieben sind. Nun werden sämtliche Verlinkungen angezeigt. Mittlerweile wurde dieses Verhalten, durch neuere Softwareversionen, ausgeschlossen.

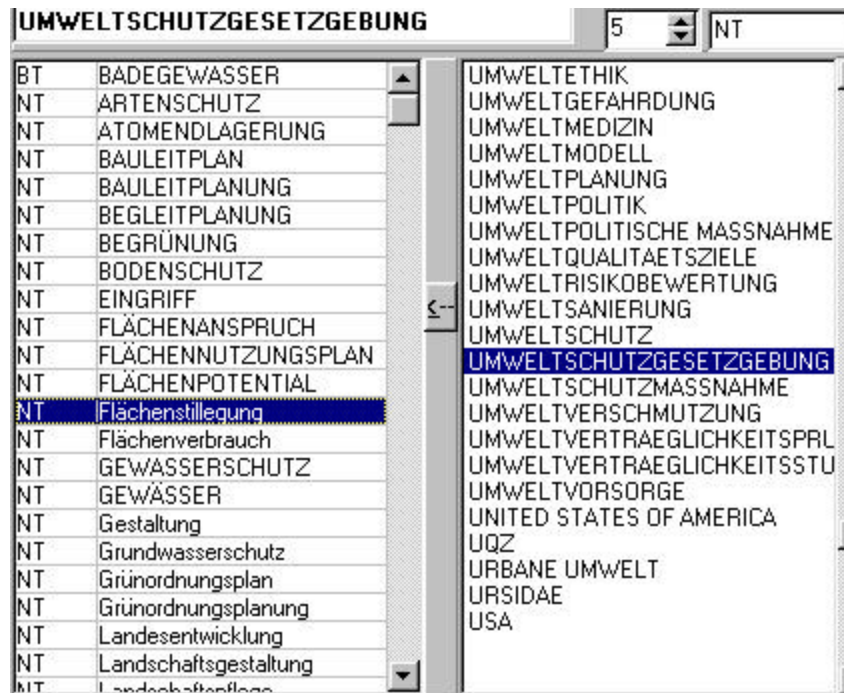


Abbildung 17 Verwandlung importierter Begriffe in thesauruseigene



Abbildung 18 Das “modify reference/note“ – Fenster

Hier wird der Begriff „Flächenstilllegung“ der ein Unterbegriff von „Umweltschutzgesetzgebung“ ist zum thesauruseigenen Term gemacht.

5.4.2 Löschen der importierten Begriffe und Sonderzeichenauflösung

Schließlich werden die nicht mehr benötigten importierten Begriffe, die nun doppelt vorkommen, gelöscht. Abbildung 19 zeigt eine Bestätigungsnachricht zur Löschung des Begriffes „Fauna“.



Abbildung 19 Bestätigungsnachricht zur Löschung des Begriffs „Fauna“

Schließlich werden die Deutschen Sonderzeichen ä, ö, ü und ß in für die Ausgabe in xml verständlichere ae, oe, ue und ss verwandelt. Abbildung 20 zeigt einen Ausschnitt aus dem fertigen Thesaurus.

ABFAELLE		0
NT	ABFALLBEHANDLUNG	ABFAELLE
NT	ABFALLBESEITIGUNG	ABFALLBEHANDLUNG
NT	ABFALLSAMMLUNG	ABFALLBESEITIGUNG
NT	ABFALLVERMEIDUNG	ABFALLSAMMLUNG
NT	ABFALLVERWERTUNG	ABFALLVERMEIDUNG
NT	ABFALLWIRTSCHAFT	ABFALLVERWERTUNG
NT	ABGAS	ABFALLWIRTSCHAFT
NT	ABWASSER	ABFLUSS
NT	ALTLAST	ABGAS
NT	ATOMENDLAGERUNG	ABGRABUNG
NT	AUSLEITUNG	ABTRAGUNG
NT	BIOLOGISCHE ABFALLBEHANDLUNG	ABWASSER
NT	DEPONIE	ABWASSERBEHANDLUNG
NT	EMISSION	ADRESSE
NT	IMMISSION	AGRAR-
NT	RECYCLING	AGRARTOURISMUS
NT	SCHADSTOFF	AKUSTIK
NT	SONDERABFALL	ALLGEMEINES
RT	SCHADSTOFFE	ALTERNATIVE
RT	UMWELTVERSCHMUTZUNG	ALTLAST
		AMPHIB
		ANALYSENVERFAHREN

Abbildung 20 fertiggestellter Thesaurus

6. Umwandlung der Thesaurusdatei in XML

Das Thesauruserstellungsprogramm „thew 32“ hat eine Schnittstelle, mit deren Hilfe die erstellte Datei in XML exportiert und somit umgewandelt werden kann. Abbildung 22 zeigt das Ergebnis.

Abbildung 21 Eingabemaske zur Umwandlung der Thesaurusdatei in XML

```

<label>KLEINLEBEWELT</label>
<parent UREF="_BIOLOGISCHE_SYSTEME_UND_PROZESSE"/>
<parent UREF="_BIOSPHERE"/>
</SVTerm>
-<SVTerm UID="_KLIMA" UIDtype="source">
<label>KLIMA</label>
<child UREF="_BIOKLIMA"/>
<child UREF="_GEMAESSIGTESKLIMA"/>
<child UREF="_KONTINENTALKLIMA"/>
<child UREF="_MEERESKLIMA"/>
<child UREF="_STADTKLIMA"/>
<child UREF="_TEMPERATUR"/>
<child UREF="_WETTER"/>
<child UREF="_WUESTENKLIMA"/>
<relatedTerm UREF="_ATMOSPHAERE"/>
</SVTerm>
-<SVTerm UID="_KOHLEKRAFTWERK" UIDtype="source">
<label>KOHLEKRAFTWERK</label>
<parent UREF="_KRAFTWERK"/>
</SVTerm>
-<SVTerm UID="_KOMMUNALPLANUNG" UIDtype="source">
<label>KOMMUNALPLANUNG</label>
<parent UREF="_KOMMUNE"/>
</SVTerm>
-<SVTerm UID="_KOMMUNE" UIDtype="source">
<label>KOMMUNE</label>
<parent UREF="_INSTITUTION"/>
<parent UREF="_ORGANISATION"/>
<parent UREF="_PLANUNG"/>
<parent UREF="_POLITIK"/>
<parent UREF="_VERWALTUNG"/>
<child UREF="_KOMMUNALPLANUNG"/>
</SVTerm>
-<SVTerm UID="_KOMPOST" UIDtype="source">
<label>KOMPOST</label>
<relatedTerm UREF="_MATERIALIEN"/>
<relatedTerm UREF="_PRODUKTE"/>
<relatedTerm UREF="_WERKSTOFFE"/>
</SVTerm>
-<SVTerm UID="_KOMPOSTIERUNG" UIDtype="source">
<label>KOMPOSTIERUNG</label>

```

Abbildung 22 Ergebnis der Umwandlung in XML Format

7. Zusammenfassung und Ausblick

In dieser Diplomarbeit machte ich zunächst eine allgemeine Einführung zur Geschichte des Thesaurus. Anschließend ging ich darauf ein, wie er im Informations- und Dokumentationswesen verwendet wird. Ferner grenzte ich den Begriff von anderen Begriffen ab.

Ich schilderte die Thesauruserstellung und die mit ihr verbundenen Probleme, wie Polysem-, Synonym-, und Homonymproblem. Ich ergänzte meine Beschreibung immer durch praktische Beispiele, auch aus meiner eigenen Thesauruserstellung, da so dem Leser die oftmals etwas komplizierte Sprache fachspezifischer Arbeiten erschlossen wird.

Ich stellte auch die beiden Umweltthesauri des Umweltbundesamtes und der Europäischen Union vor, da insbesondere letzterer zur Grundlage meiner Bearbeitung des Firmenthesaurus wurde. Dann ging ich noch auf die unterschiedliche Thesaurussoftware ein.

Nachdem ich die Ausgangslage in meiner Firmenbibliothek dargestellt hatte, beschrieb ich das Vorgehen, um das Ziel meiner Arbeit, einen Thesaurus für die Firma zu erstellen, zu erreichen.

Ich importierte alle Schlagwörter in die Thesaurussoftware und verlinkte sie, so dass sie der Verlinkung im GEMET entsprachen. Schließlich gab ich das Ergebnis in einer XML-Datei aus.

Rückblickend lässt sich sagen, dass ich meine Diplomarbeit als gute Grundlage für den weiteren Ausbau des Thesaurus verwenden kann. Es ist nützlich, nun eine geordnete Schlagwortsammlung nach GEMET zu haben, da am Anfang jedes Schlagwort ja nur in alphabetischer Reihenfolge vorhanden war. Als weiteres Projekt kann zum Beispiel der Thesaurus an eine Bücherdatenbank angefügt werden. So kann er deren Schlagwörter in hierarchischen Beziehungen anzeigen.

Abschließend möchte ich noch bemerken, dass mir das Erstellen des Thesaurus großen Gewinn brachte, da ich tiefer in die Anordnung der Deskriptoren als auch in die Umweltthematik eingeführt wurde.

Ich hoffe, dass meine Arbeit in der Firmenbibliothek zur häufigen Anwendung kommen wird.

Glossar

ahierarchisch: nicht hierarchisch

alternieren: abwechseln (Herder, 1971)

a posteriori: erfahrungsgemäss Gegensatz: a priori (Herder, 1971)

a priori: von vornherein, aus Vernunftgründen. Gegensatz: a posteriori (Herder, 1971)

denotativ: nur den begrifflichen Inhalt betreffend, ohne Nebenbedeutungen und Nuancen, z.B. Krebs (Tier) Gegensatz: konnotativ (Teachsam, 2003)

konnotativ: die sprachlichen Nebenbedeutungen und Bedeutungsnuancen betreffend. z.B. Krebs (Krankheit und der Schrecken, welcher von ihr ausgeht) (Teachsam, 2003)

kontradiktorisch: sich gegenseitig aufhebenden (Herder, 1971)

Postkoordination: „ist das Indexierungsprinzip, nachdem Kombinationen von Begriffen *während der Suche* durch Kombination von Indextermini hergestellt werden.“

(Neveling, 1975 S. 67)

Präkoordination : „ist das Indexierungsprinzip, nachdem Kombinationen von Begriffen *während des Indexierens* hergestellt werden.“

(Neveling, 1975 S. 68)

reziprok: wechsel-, gegenseitig (Herder, 1971)

Quellenangaben

(Aitchison und Gilchrist, 1987): Aitchison, Jean, Gilchrist, Alan: Thesaurus - construction: a practical manual.- 2nd ed., London: Aslib, 1987

(Batschi, 1994): Wolf-Dieter: Environmental Thesaurus and Classification of the Umweltbundesamt (Federal Environmental Agency), Berlin, 1994 (S. 57-62) in: Stancikova, Pavla; Dahlberg, Ingetraut (Hrsg.): Environmental Knowledge Organization and Information Management.- Frankfurt/Main: Indeks Verl., 1994

(Burkart, 1997): Margarete: Thesaurus (S. 160-179) in: Marianne Buder et al. (Hrsg.) Grundlagen der praktischen Information und Dokumentation: ein Handbuch zur Einführung in die fachliche Informationsarbeit.- 4. völlig neu gefasste Aufl.- München: Saur, 1997

(Craven, 2003): Tim: Thesaurussoftware <http://publish.uwo.ca/~craven/freeware.htm> zuletzt besucht am 02.09.03

(ESI, 2001) ESI (Environnement Scientifique Intégré) : Thesaurus Definition

<http://www.esi.umontreal.ca/~saadanil/Acsi2/tsld004.htm>

zuletzt besucht am 29.08.2003

(Herder, 1971): Herders Fremdwörterbuch zus. gest. v. Eugen F. Kuri. Freiburg: Herder, 1971

(GEMET, 2001): European Topic Centre on Catalogue of data sources (ETC/CDS) ; European Environment Agency : Preface to the 2001 version of GEMET

http://eea.eionet.eu.int:8980/irc/Download/kfeYA6JGmUGsG67_BioqznZp5sVu1GjI/d-4yHh4tk6f4g-4/Info2001.pdf zuletzt besucht am 29.08.03

(Laisiepen et al. 1972): Grundlagen der praktischen Information und Dokumentation. München, 1972. In: Wersig, Gernot: Thesaurus-Leitfaden: eine Einführung in das Thesaurusprinzip in Theorie und Praxis. - München: Verl. Dokumentation Saur, 1978, S. 22

(Legat, 1999): Rudolf, Stallbaumer, Hermann: Erstellung, Wartung und Visualisierung multilingualer Thesauri mittels "THESmain" und "THESshow" aus: Management von Umweltinformationen in vernetzten Umgebungen. 2. Workshop des Arbeitskreises "[Hypermedia im Umweltschutz](#)". S. 116. Umwelt-Informatik aktuell, Band 21. Metropolis Verlag, Nürnberg, März 1999. http://www.cedar.at/wgr_home/pub/thesmain-masterdoc.html#5.1 zuletzt besucht am 05.09.03

(Lexikon, 1968): Der Neue Herder in sechs Bänden, 1968 Begriffe: „Thesaurus“ und „Schatzhäuser“

(Manecke, 1997): Hans-Jürgen: Klassifikation (S. 141-159) in: Marianne Buder (Hrsg.) et al. Grundlagen der praktischen Information und Dokumentation: ein Handbuch zur Einführung in die fachliche Informationsarbeit.- 4. völlig neu gefasste Aufl.- München: Saur, 1997

(Müller, 2003): Reinhard: Realisierung einer ACCESS Datenbank in einer Firmenbibliothek zu Stuttgart: Große Studienarbeit im Wahlpflichtfach Informationstechnik des Studiengangs Informationswirtschaft an der Fachhochschule Stuttgart – Hochschule der Medien, 2003

(Neveling, 1975): Ulrich; Wersig, Gernot: Terminologie der Information und Dokumentation.- München: Verl. Dokumentation, 1975

(Nohr, 1993): Holger: Ordnung schaffen: Thesauri in integrierter Dokumentationssoftware für den PC. in: Zeitschrift cogito 9 (1993) 3, S. 28-30

<http://www.iuk.hdm-stuttgart.de/nohr/publ/Thesaurussoftware.pdf> zuletzt besucht am 29.08.2003

(Nohr, 2000): Holger: Thesaurus zum menschlichen Skelett. Erstellt von Jasmin Wendt und Gundula Marschke im Sommersemester 00 im Fach „Inhaltliche Erschließung“ im Studiengang „Informationsdesign“

<http://www.iuk.hdm-stuttgart.de/nohr/studarb/Skelett/Eingangsseite.htm> zuletzt besucht am 29.08.2003

Thesaurus zum Thema Fahrrad. Erstellt von Katja Richter und Matthias Kissler im Sommersemester 00 im Fach „Inhaltliche Erschließung“ im Studiengang Informationsdesign. <http://www.iuk.hdm-stuttgart.de/nohr/studarb/Fahrrad/Thesaurus.html> zuletzt besucht am 29.08.2003

(Reusch, 2001): Peter J.A. (Fachhochschule Dortmund) : <http://www.wirtschaft.fh-dortmund.de/~reusch/prothe1.html>

zuletzt besucht am 29.08.2003

(Riekert, 1996): Wolf-Fritz: Thesaurusgestützter Zugriff zu Umweltinformation in einem netzübergreifenden Hypermediasystem.- Ulm: Forschungsinstitut für anwendungsorientierte Wissensverarbeitung (FAW), 1996

(Soergel, 1971): Dagobert: Dokumentation und Organisation des Wissens: Versuch einer methodischen und theoretischen Grundlegung am Beispiel der Sozialwissenschaften.- Berlin: Duncker & Humbolt, 1971

(Spree, 2003): Ulrike: (Hochschule für angewandte Wissenschaften Hamburg, Fachbereich Bibliothek und Information) <http://www.bui.fh-hamburg.de/pers/ulrike.spree/grundthes/grundthes.htm> zuletzt besucht am 29.08.2003

(Teachsam, 2003) Teachsam: Lehren und lernen online, 2003.
http://www.teachsam.de/glossar_deu.htm zuletzt besucht am 29.08.03

(Umstätter, 1999): Walther: Die vernachlässigten Aspekte des Thesaurus: dokumentarische, pragmatische, semantische und syntaktische Einblicke. in: Schwarz, I./ Umstätter, W. : nfd Information – Wissenschaft und Praxis 50 (4) S. 197-203 (1999)
<http://www.ib.hu-berlin.de/~wumsta/infopub/pub1996f/pub110.html> zuletzt besucht am 29.08.2003

(Umstätter, 2003): Walther: dies ist die Diskussionsgrundlage für einen semiotischen Thesaurus der Bibliothekswissenschaft. <http://www.ib.hu-berlin.de/%7Ewumsta/wistru/definitions/hierdef01b.pdf> zuletzt besucht am 29.08.03

(Wersig, 1978): Gernot: Thesaurus - Leitfaden: eine Einführung in das Thesaurusprinzip in Theorie und Praxis. - München: Verl. Dokumentation Saur, 1978

(Wersig, 1985): Gernot: Thesaurus Leitfaden. 2. erg. Aufl.. München, 1985 in:: Burkart, Margarete: Thesaurus (S. 160-179) in: Marianne Buder et al.(Hrsg.) Grundlagen der praktischen Information und Dokumentation: ein Handbuch zur Einführung in die fachliche Informationsarbeit.- 4. völlig neu gefasste Aufl.- München: Saur, 1997

(Wikipedia, 2003): Wikipedia, die freie Enzyklopädie, 2003. Stichwort „Thesaurus“
<http://de.wikipedia.org/wiki/Thesaurus> letzter Zugriff am 29.08.2003

Erklärung

Hiermit erkläre ich, dass ich die vorliegende Diplomarbeit selbständig angefertigt habe. Es wurden nur die in der Arbeit ausdrücklich benannten Quellen und Hilfsmittel benutzt. Wörtlich oder sinngemäß übernommenes Gedankengut habe ich als solches kenntlich gemacht.

Stuttgart, 24.09.2003

Ort, Datum

Reinhard Müller

Unterschrift