



Anreicherung des Zugangsvokabulars zur DDC durch multilinguale Schlagwörter

Das Projekt CrissCross

vorge stellt am 08.03.2007 von

Katrin Müller und Tina Mengel, FH Köln



Projektdaten

Projekträger:

Deutsche Forschungsgemeinschaft (DFG)

Projektpartner:

Deutsche Nationalbibliothek, Fachhochschule Köln

Projektausführende Personen an der FH Köln:

Anne Betz, Jessica Hubrich, Tina Mengel, Katrin Müller

Dauer des Projekts: 2006 – 2008

Internetseite:

<http://www.d-nb.de/wir/projekte/crisscross.htm>

Projektleitung in DNB:

Helga Karg

h.karg@d-nb.de

Projektleitung FH Köln:

Prof. Winfried Gödert

winfried.goedert@fh-koeln.de





Inhalt dieses Vortrags:

- Grundproblematiken und CrissCross-Vorgeschichte
- CrissCross – Ziele und Maßnahmen
- Beispiele aus dem Mapping-Alltag – Determiniertheit und Methodik
- Relationenerweiterungen und Retrievallösungen



Grundproblematiken

Ausgangslage und Gewolltes

Es existieren Dokumentationssprachen...

- a) in unterschiedlicher Form
 - (Struktur)heterogenität der Systeme
(Thesauri, Klassifikationssysteme)

- b) in unterschiedlichen Landessprachen
 - Suche = sprachabhängig

- c) deren Struktur und Inhalt an die lokalen Gegebenheiten angepasst sind
 - Lokalisierung nach kulturellen, politischen, rechtlichen etc. Wirklichkeitskonstrukten

Die Suche nach erschlossenen Dokumenten...

setzt Kenntnisse über die jeweils verwendeten Erschließungssysteme voraus

ist nur in Beständen erfolgreich, für deren Erschließung dieselbe Dokumentations-
sprache verwendet wird

in verbaler Form ist an die Landessprache(n)
gebunden, in der auch die Dokumenta-
tionssprache vorliegt





Wissenssuche heute sollte...

einfacher sein

- Suche in vielen, miteinander verbundenen Katalogen
- Unabhängigkeit der Suche von den jeweiligen Erschließungssystemen
- benutzerorientierte Sucheinstiege

international sein

- Suche in internationalen Beständen unter Verwendung der eigenen Landessprache als Suchsprache
- Lokalisierung der Systeme sollte für das Retrieval keine Hindernisse darstellen

zukunftsorientiert sein

- Austauschformate; Möglichkeit für Erweiterungen und Weiterentwicklung
- Bereitsein für die Anpassung an neue Technologien (wie z. B. semantische)
- Kooperation mit Bibliotheken anderer Länder





Vorangegangenes und Gegebenes

MACS (Multilingual Access to Subject Headings)

1999 von CENL initiiert

Beteiligt: DNB, SLB, BNF, BLL

Aufgabenstellung/ Zielsetzung:
Ermöglichung der multilingualen Suche nach
verbal erschlossenen Dokumenten

Realisierung:
Mittels Zusammenführung von Schlagwörtern
der SWD, LCSH und RAMEAU



Vorgehensweise:

SWW aus den Bereichen Sport und Theater sowie die 500 am häufigsten von der BNF benutzten SWW und deren Äquivalente in SWD und LCSH

Paradigmen zur Verknüpfung der SWW:

- Volläquivalenzen
children = enfant = Kind
- Symmetrie der Verknüpfungsrelation
- child actors = enfants acteurs = Kind AND
Schauspieler => ???
- keine Festlegung einer Leitsprache

Offene Fragen:

Anbindung zusätzlicher Sprachen?

Lokalisierung von Begriffen?

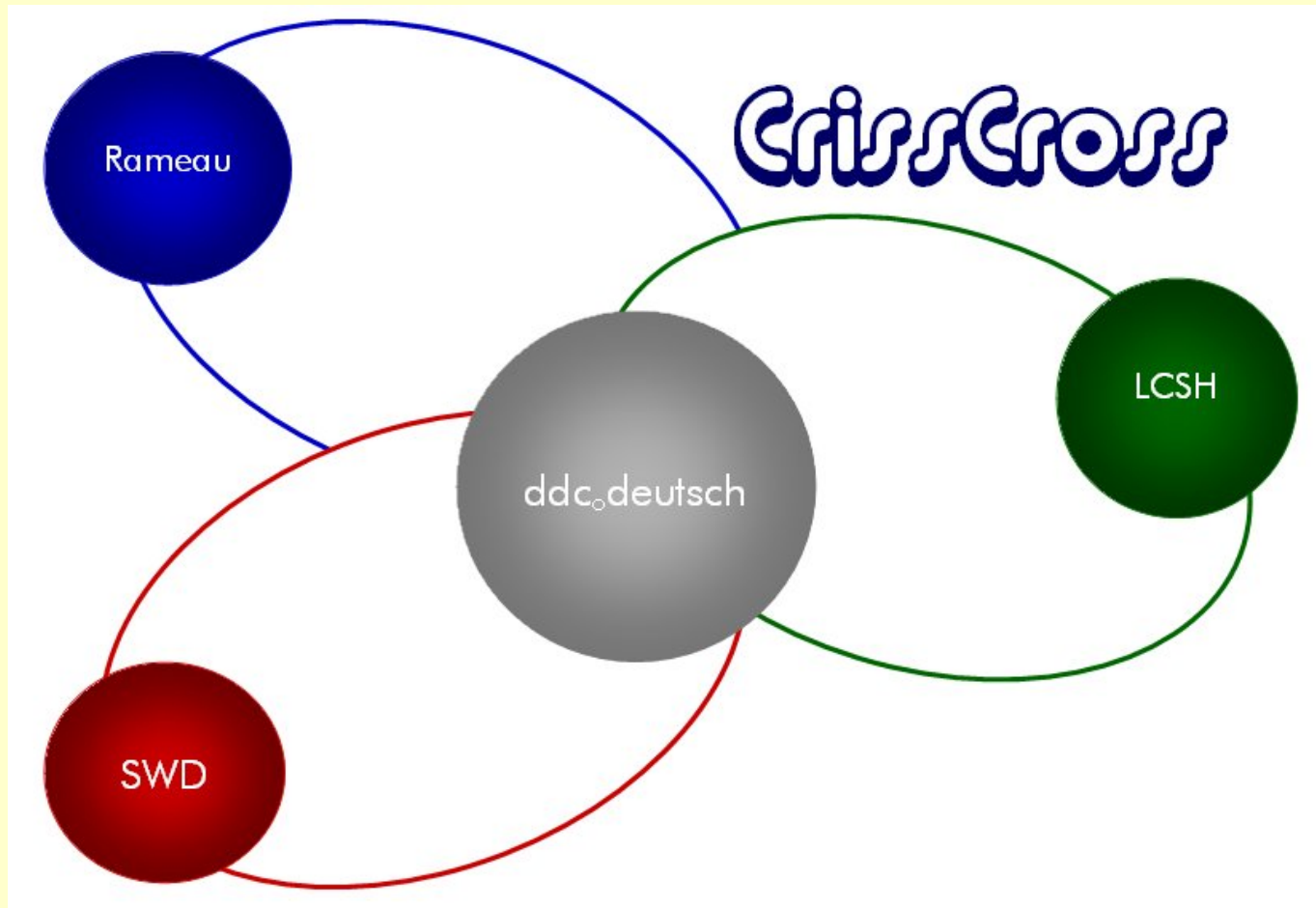
Umgang mit Strukturheterogenität?





Projekt CrissCross

Ein neuer Ansatz



Projektziel

Schaffung eines multilingualen, thesaurusbasierten und benutzergerechten Recherchevokabulars zu heterogen erschlossenen Dokumenten für die Nutzer des deutschen Bibliothekswesens durch Anreicherung des DDC-Registervokabulars



Maßnahmen

- Integration der im deutschsprachigen Raum und weltweit am häufigsten genutzten Dokumentations-sprachen (SWD, LSCH, Rameau, DDC)
- DDC als Backbone → Erweiterung von MACS um klassifikatorische Komponente
Vorteile DDC:
 - numerisch, daher sprachunabhängig
 - hierarchisches Browsing möglich
- Thesauri werden an dieses Backbone angeschlossen
 - Zusammenführung von Klassifikation und Thesauri
 - Erweiterung des Zugangsvokabulars zu DDC-erschlossenen Dokumenten



Erstes Arbeitspaket:

Verbindung der DDC Deutsch mit der Schlagwortnormdatei (SWD):

SWD-Sachschlagwörter werden einer oder mehreren Notation(en) der DDC zugeordnet und es wird die Bindungsstärke des SW zur DDC-Klasse festgelegt

Die Verlinkung (das Mapping) geschieht in der WinIBW (Pica/ILTIS-Arbeitsumgebung) direkt im SW-Datensatz

Zur Recherche dienen: MelvilClass, ClassWeb, OCLC Connexion, WorldCat, Bibliothekskataloge wie der KVK, das WWW



Zweites Arbeitspaket:

Anbindung 70.000 bestehender Rameau-LCSH-Tupel an das mit DDC gemappte SW; Verlinkung der Tupel in der SWD

Vorschläge für die Darstellung der Begriffsbeziehungen mit Hilfe semantischer Technologien; Skizzierung von Anwendungsmöglichkeiten

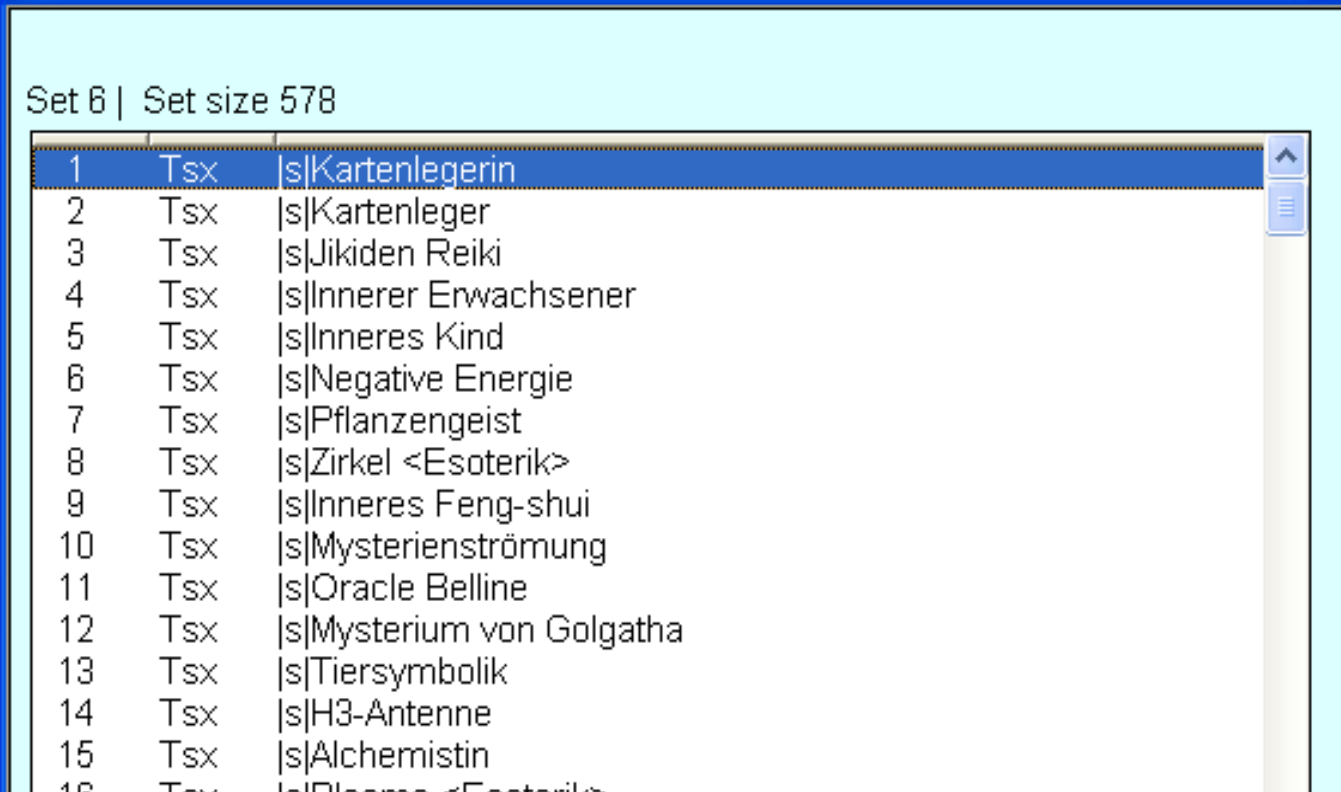




Beispiele aus dem Mapping-Alltag

Vorgehensweise beim Mapping SWD-DDC

Aufteilung der Arbeitsbereiche nach SWD-Sachgruppen
(über Systematiknummern), z. B. 5.7 Esoterik



Set 6 | Set size 578

1	Tsx	s Kartenlegerin
2	Tsx	s Kartenleger
3	Tsx	s Jikiden Reiki
4	Tsx	s Innerer Erwachsener
5	Tsx	s Inneres Kind
6	Tsx	s Negative Energie
7	Tsx	s Pflanzengeist
8	Tsx	s Zirkel <Esoterik>
9	Tsx	s Inneres Feng-shui
10	Tsx	s Mysterienströmung
11	Tsx	s Oracle Belline
12	Tsx	s Mysterium von Golgatha
13	Tsx	s Tiersymbolik
14	Tsx	s H3-Antenne
15	Tsx	s Alchemistin
16	Tsx	s Plasma <Esoterik>



Streuungen in der SWD

SWD-Sachgruppe		Verknüpfte SW
00	Unspezifische Allgemeinwörter	1
1	Allgemeines, Interdisziplinäre Allgemeinwörter	7
2	Schrift, Buch, Presse	15
3	Religion	193
4	Philosophie	83
5	Psychologie, Esoterik	2532
6	Kultur, Erziehung, Bildung, Wissenschaft	1079
7	Recht, Verwaltung	55
8	Politik, Militär	22
9	Soziologie, Gesellschaft, Arbeit, Sozialgeschichte	430
10	Wirtschaft, Verkehr, Umweltschutz, Raumordnung	98
11	Sprache	50
12	Literatur	21
13	Bildende Kunst, Photographie	18
14	Musik	23
15	Theater, Tanz, Film, Rundfunk	11
16	Geschichte	7
17	Volkskunde, Völkerkunde	38
18	Natur, Naturwissenschaften allg.	6
19	Geowissenschaften	23

20	Astronomie, Weltraumforschung	35
21	Physik	34
22	Chemie	112
23	Allg. Biologie, Mikrobiologie	1127
24	Botanik	49
25	Zoologie	371
26	Anthropologie	17
27	Medizin	6267
28	Mathematik	47
29	Stochastik, Operations Research	20
30	Informatik, Datenverarbeitung	120
31	Technik	173
32	Landwirtschaft, Gartenbau	321
33	Hauswirtschaft, Körperpflege, Mode, Kleidung	49
34	Sport	257
35	Spiel, Unterhaltung	219
36	Basteln, Handarbeiten, Heimwerken	0
	Gesamtzahl der vollständig bearbeiteten Sachschlagwörter	13930
	<i>Sachschlagwörter ohne Determiniertheit (Vorschläge der DNB)</i>	1825
	Gesamtzahl der verlinkten Sachschlagwörter	15755



Streuungen in der DDC

DDC-Hauptklasse		Vergebene Notationen
000	Informatik, Informationswissenschaft, allgemeine Werke	523
100	Philosophie und Psychologie	1671
200	Religion	331
300	Sozialwissenschaften	2448
400	Sprache	105
500	Naturwissenschaften und Mathematik	3513
600	Technik, Medizin, angewandte Wissenschaften	7593
700	Künste und Unterhaltung	610
800	Literatur	34
900	Geschichte und Geografie	25
	Gesamtzahl vergebener Notationen	16853

Aufrufen des SW-Datensatzes

Set 3 | Set size 12 | Record 4 | PPN 964148315 | Format DA

Eingabe: 1250:03-04-02 Änderung: 0832:21-02-07 16:21:58 Status:
0832:21-02-07

005 Ts

012 /xxxw

021 4680522-9

800 **|s|Kipper-Wahrsagekarten**

808 |a|-Graf,Eckhard : Lexikon d. Tarot

808 |b|Kartenorakel von Susanne Kipper, entstanden um 1890

810 5.7

830 |s|Kipperkarten

830 |s|Kipper-Karten



Relevante Klasse in MelvilClass

Tina Mengel
FH Köln - CrissCross

MelvilClass
The dot in 025

Klassen Suche     Hilfe Logout ◀ ▶ Suche:

Klassen

Notation	Bezeichnung
	DDC-Übersicht
100	Philosophie & Psychologie
130	Parapsychologie & Okkultismus
133	Einzelne Themen der Parapsychologie und des Okkultismus
133.3	Divinatorische Künste
	Wahrsagen mit Kristallen und Steinen; Rutengehen;
133.32	Wahrsagen mit Karten, Teeblättern und Kaffeesatz, durch Orakel und Sibyllen
	Wahrsagen mit Karten, Teeblättern und Kaffeesatz, durch Orakel und Sibyllen
133.324	
133.3242	Wahrsagen mit Karten (Kartomantie)
 133.3242092	Kartenleger
133.32424	Wahrsagen mit Tarot

Anmerkungen, Hinweise und Verweisungen

Registereinträge

- [Karten — Divination](#)
- [Kartenlegen](#)
- [Kartomantie](#)



DDC-Klasse mit ergänzten SWW

Set 3 | Set size 12 | Record 4 | PPN 964148315 | Format DA

Eingabe: 1250:03-04-02 Änderung: 0832:21-02-07 16:23:07 Status:
0832:21-02-07

005 Tsx

012 /xxw

021 4680522-9

800 **|s|Kipper-Wahrsagekarten**

808 |x|me

808 |a|-Graf,Eckhard : Lexikon d. Tarot

808 |b|Kartenorakel von Susanne Kipper, entstanden um 1890

810 5.7

816 **133.3242#2#**

830 |s|Kipperkarten

830 |s|Kipper-Karten



Der Grad der Determiniertheit

Die Bindungsstärke zwischen einem SW und einer DDC-Klasse wird durch die Vergabe der Ziffern 3, 2, und 1 festgelegt.

005 Tsx

012 /ooW

021 4003657-1

800 |s|Aura

808 |x|me

808 |a|M 2., Dorsch

810 5.6;5.7

816 133.892#3#

816 612.0142#2#

Die Ziffern

signalisieren die Übereinstimmung des SW mit der DDC-Klasse (3 = große, 2 = mittlere, 1 = geringe Übereinstimmung)

sollen später für das Ranking der Trefferliste in einem bibliografischen Webrecherchedienst ein ausschlaggebender Faktor sein.



DDC-Klasse mit ergänzten SWW

Tina Mengel
FH Köln - CrissCross

MelvilC
The dot in 025

Klassen Suche     Hilfe Logout ◀ ▶ Suche:

Klassen

Notation	Bezeichnung
	DDC-Übersicht
100	Philosophie & Psychologie
130	Parapsychologie & Okkultismus
133	Einzelne Themen der Parapsychologie und des Okkultismus
133.3	Divinatorische Künste
	Wahrsagen mit Kristallen und Steinen; Rutengehen;
133.32	Wahrsagen mit Karten, Teeblättern und Kaffeesatz, durch Orakel und Sibyllen
	Wahrsagen mit Karten, Teeblättern und Kaffeesatz, durch Orakel und Sibyllen
133.324	
133.3242	Wahrsagen mit Karten (Kartomantie)
 133.3242092	Kartenleger
133.32424	Wahrsagen mit Tarot

Anmerkungen, Hinweise und Verweisungen

Registereinträge

- [Karten — Divination](#)
- [Kartenlegen](#)
- [Kartomantie](#)

SWD-Schlagwörter

- [|s|Biedermeier-Aufschlaqkarten](#)
- [|s|Engelkarten](#)
- [|s|Karma-Karten](#)
- [|s|Kartenlegen](#)
- [|s|Keltisches Baumorakel](#)
- [|s|Kipper-Wahrsagekarten](#)
- [|s|Lenormand-Karten](#)
- [|s|Mô <Kartenspiel>](#)
- [|s|Seni-Horoskop](#)
- [|s|Symbolon <Spiel>](#)
- [|s|Zeitqualität](#)
- [|s|Zigeuner-Wahrsagekarten](#)



Ansetzungsform: **|s|Kipper-Wahrsagekarten****|s|Kipper-Wahrsagekarten**

IDN 964148315

BF |s|Kipperkarten

BF |s|Kipper-Karten

DDC [133.3242](#)

Beispiel für Strukturheterogenität:
SWD-OB und UB = beide 133.3242

Ansetzungsform: **|s|Kartenlegen**UB1 [|s|Biedermeier-Aufschlagkarten](#)UB1 [|s|Lenormand-Karten](#)UB1 [|s|Oracle Belline](#)☐ UB1 [|s|Tarot](#) ▲ ▶UB2 [|s|Akron-Tarot](#)UB2 [|s|Antichi Tarocchi Esoterici](#)UB2 [|s|Astrologisches Mandala-Tarot](#)UB2 [|s|Cat-people-Tarot](#)**|s|Kartenlegen**

IDN 040297950

BF |s|Kartenschlagen

UB |s|Biedermeier-Aufschlagkarten

UB |s|Lenormand-Karten

UB |s|Oracle Belline

UB |s|Tarot

DDC [133.3242](#)

Sachgruppe Esoterik 5.7

Merkmale:

- o SW oft **allgemeiner** als bestehende DDC-Klasse, z. B. SW Esoterik, Okkultismus, Spuk
- o **Große Streuung** innerhalb der DDC, z. B. SW Übernatürliches Wesen → **133.4#1#** Dämonologie und Hexerei **133.9#1#** Spiritismus, **202.11#2#** [...], Gottheiten, **398.21#2#** Erzählungen [...] paranatürliche Wesen [...]; z. B. SW Gedankenkraft → **133.43#1#** Magie und Hexerei, **153.42#1#** Denken und Nachdenken, **153.3#1#** Imagination, Vorstellungskraft, Kreativität
- o **Auslagerung von SWW aus der SG**, weil in der DDC anderer Bezug, z. B. SW Earthwalk → **204.4#1#** Religiöses Leben und religiöse Praxis, **201.44#1#** Schamanismus, **796.51#1#** Wandern; z. B. SW Freimaurerei → DDC **366.1#3#** Freimaurerei
- o **Wenig** neu erstellte Notationen
- o **Viele** fehlende Relationen (über 180 von 578 SWW)
- o **Wenig** vorhandene OBs (= SWW der SG, die selbst OB sind; ca. 36)

Sachgruppe Medizin 27

Merkmale:

- o SW oft **spezifischer** als bestehende DDC-Klasse
- o **Hohe Anzahl** neu erstellter DDC-Notationen mithilfe der Anhängetafeln in der Medizin (s. 616.1-616.9 Einzelne Krankheiten) und den Biowissenschaften (s. 573 Einzelne physiologische Systeme bei Tieren, regionäre Histologie und Physiologie bei Tieren)

Beispiele:

573.17 Pumpmechanismen in der Biologie
→ SW Herz #3#

573.17 + 27 Muskeln des Systems
→ SW Herzmuskel #3#

573.1727 + 36 Zellbiologie
→ SW Herzmuskelzelle #3#

616.14 + 06 Krankheiten der Venen und Kapillaren—Therapie
→ SW Kompressionsbehandlung #2#

- o **Viele** fehlende Relationen

Nutzungsmöglichkeiten der Mappings in einer Retrievalumgebung



- Sachliche Suche mit (multilingualem) verbalem Zugang wird verbessert
- Auswahlmöglichkeit von Kontext, in dem das SW stehen soll
- DDC-Hierarchie kann zum Ausweiten/ Einschränken der Suche genutzt werden
- Nutzung der Erschließungsdaten aus Fremddaten
- Nutzen der Determiniertheit für das Trefferranking

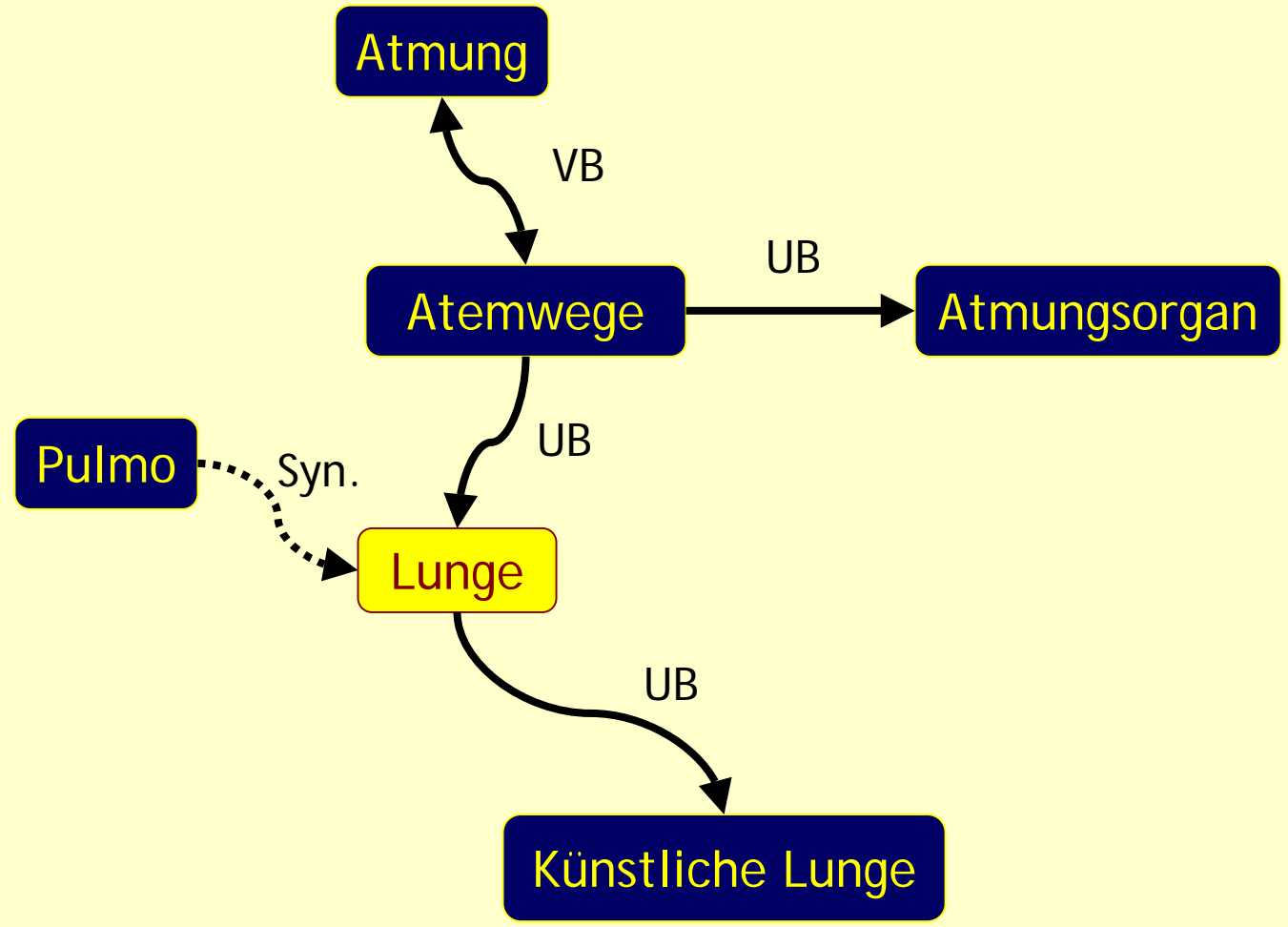
616.128 Arrhythmie

SW Arrhythmie#3#

SW Kammerflimmern#2#

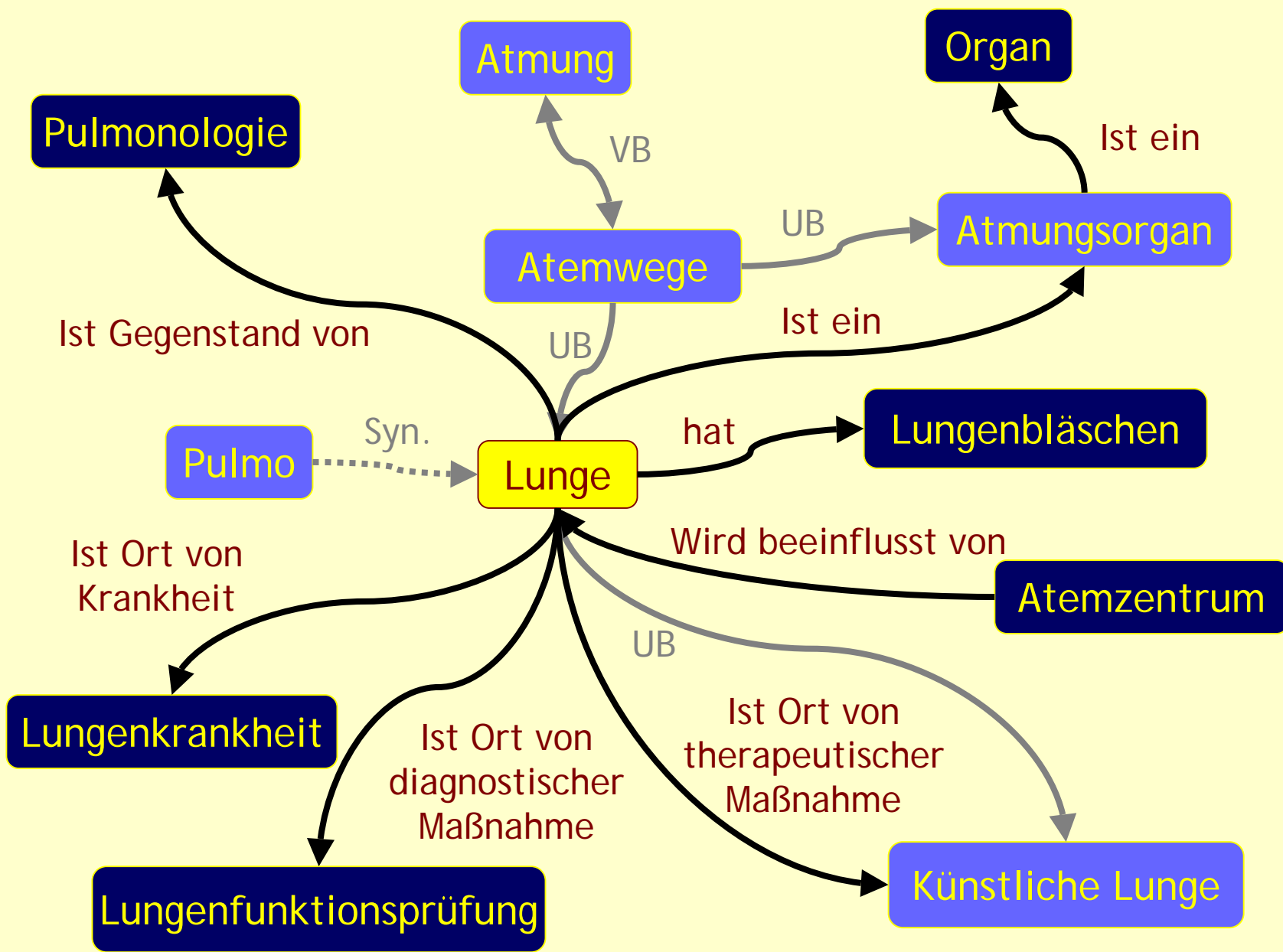
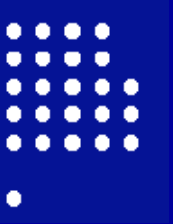
SW Alternans#1#





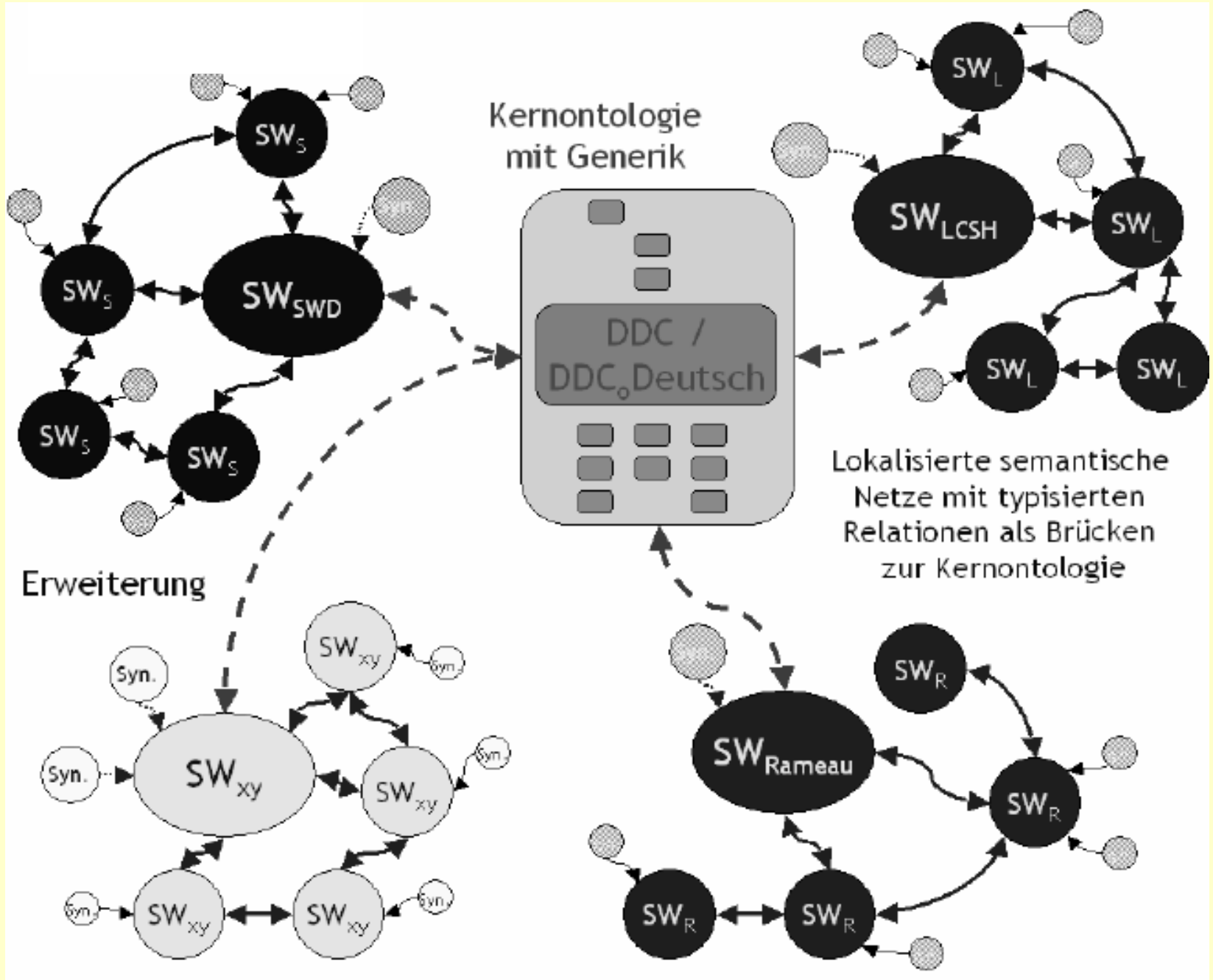
Das Schlagwort *Lunge* in der *SWD*





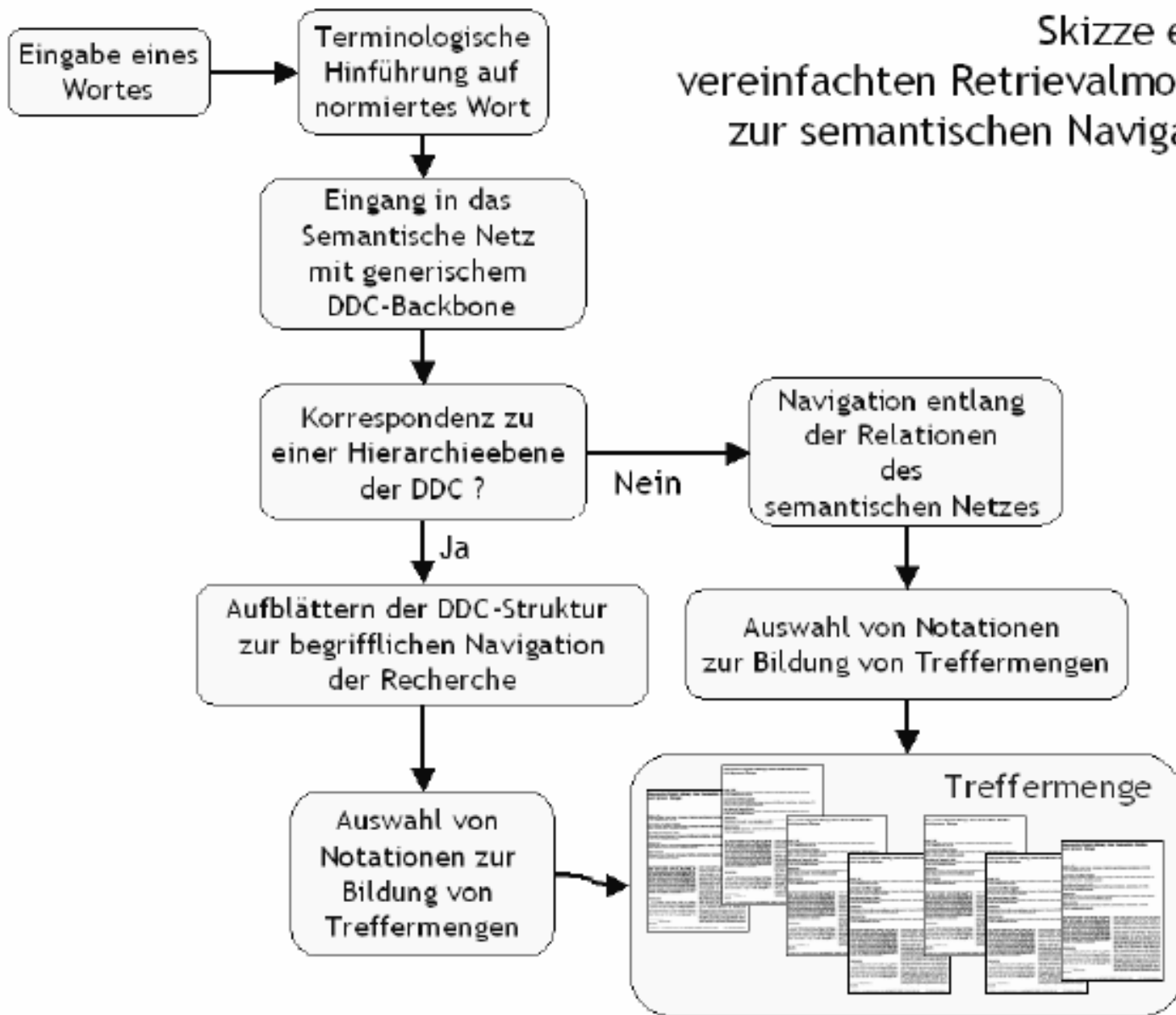
Beispiel für erweiterte Relationierung





Integration in Retrievalmodelle

Skizze eines vereinfachten Retrievalmodells zur semantischen Navigation





**Vielen Dank für Ihre
Aufmerksamkeit!**

Projektdaten

Projektträger:

Deutsche Forschungsgemeinschaft (DFG)

Projektpartner:

Deutsche Nationalbibliothek, Fachhochschule Köln

Projektausführende Personen an der FH Köln:

Anne Betz, Jessica Hubrich, Tina Mengel, Katrin Müller

Dauer des Projekts: 2006 – 2008

Internetseite:

<http://www.d-nb.de/wir/projekte/crisscross.htm>

Projektleitung in DNB:

Helga Karg

h.karg@d-nb.de

Projektleitung FH Köln:

Prof. Winfried Gödert

winfried.goedert@fh-koeln.de

