

Einsatzmöglichkeiten von Crosskonkordanzen in vascoda

*Philipp Mayr, Anne-Kathrin Walter
Informationszentrum Sozialwissenschaften, Bonn*

12. Jahrestagung der IuK-Initiative der wissenschaftlichen
Fachgesellschaften in Deutschland

28. September 2006 in Göttingen



Agenda

- Einführung: semantische Heterogenität und Heterogenitätsbehandlung
- Einsatz von Crosskonkordanzen
- Projekte mit Crosskonkordanzen
 - Übersicht fertige Crosskonkordanzen
- Verfügbarkeit
 - Heterogenitätsservice
- Ausblick



Einführung – Semantische Heterogenität

Ziel der Informationsanbieter:

- Integrierte Recherche über mehrere Informationsangebote
- Bündelung von Informationstypen
- Interdisziplinäre Suche

Aber:

Unterschiedliche Erschließungssysteme der Informationsanbieter

→ semantische Heterogenität



Heterogenitätsbehandlung

Verschiedene Ansätze, u.a.:

- Statistische Verfahren
- Intellektuelle Verfahren: **Crosskonkordanzen**

Crosskonkordanzen:

*Gerichtete, relevanzbewertete Relationen
zwischen Termen zweier Thesauri oder
Klassifikationen*

Was sind Crosskonkordanzen?

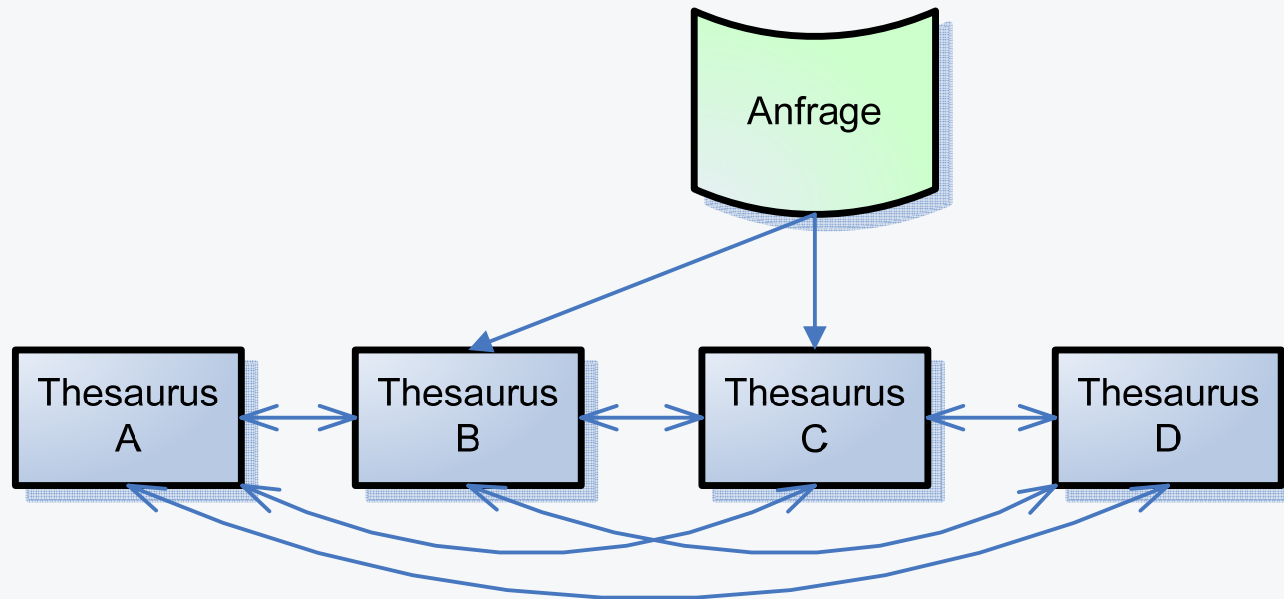
Relationen:

Ausgangsterm	Relation	Relevanz	Zielterm
Schifffahrtsrecht	=+o (Äquivalenz)	m	Schifffahrt + Verkehrsrecht
Schifffahrtsrecht	<o (Oberbegriff)	g	Verkehrsrecht
Schifffahrtsrecht	^o (verwandt)	m	Seerecht
Regionalstruktur	> (Unterbegriff)	m	Regionale Wirtschaftsstruktur

Relevanzen bezogen auf die Treffermenge

1:1- oder 1:n-Verknüpfungen

Einsatz von Crosskonkordanzen

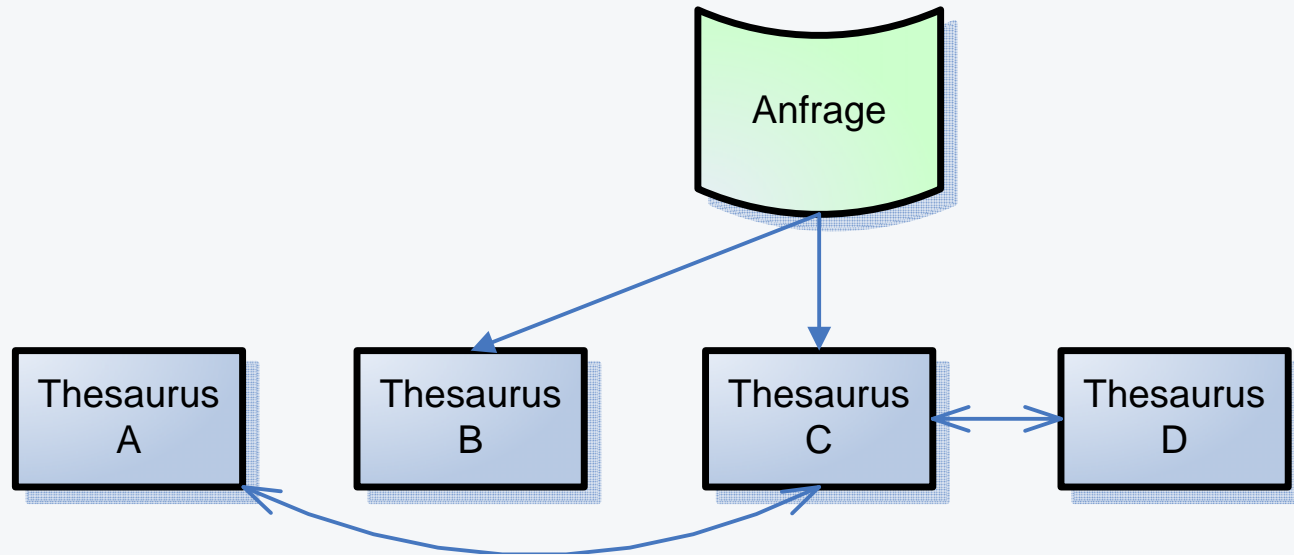


Anfrage kann in Thesaurus B und in Thesaurus C abgebildet werden

- Welche Crosskonkordanzen werden angewendet?
- Entstehung von komplexen Suchanfragen

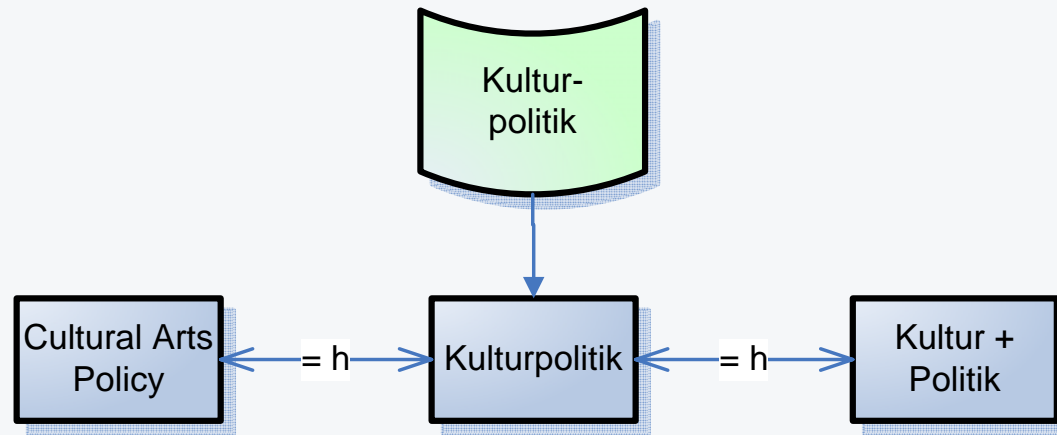
(Praxis: keine vollständige Vermaschung der Vokabulare)

Einsatz von Crosskonkordanzen



Wahl eines Ausgangsthesaurus

Einsatz von Crosskonkordanzen: Ausgangsthesaurus

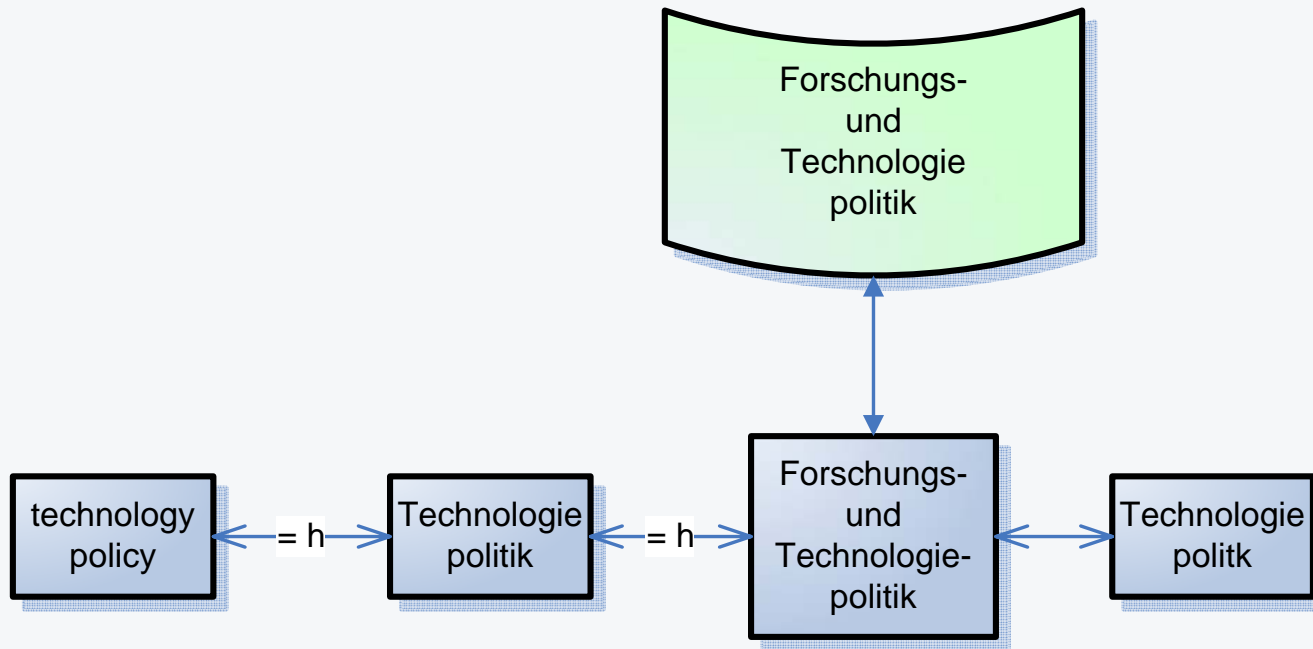


Möglichkeiten:

- Term wurde aus graphischer Oberfläche ausgewählt
- Einschränkung der Suche auf bestimmte Fächer
- Heuristiken, z.B. Trefferanzahl pro Term, Anzahl der Crosskonkordanzen des Vokabulars,...

Strategien zur Wahl des Ausgangsthesaurus sind noch zu testen!

Einsatz von CK: indirekte Transformationen



Indirekte Transformationen, um Aufwand der Crosskordanzerstellung gering zu halten.



Projekte mit Crosskonkordanzen

infoconnex:

Informationsverbund Bildung—Sozialwissenschaften—Psychologie
Projektlaufzeit Juli 2001 bis Mai 2005

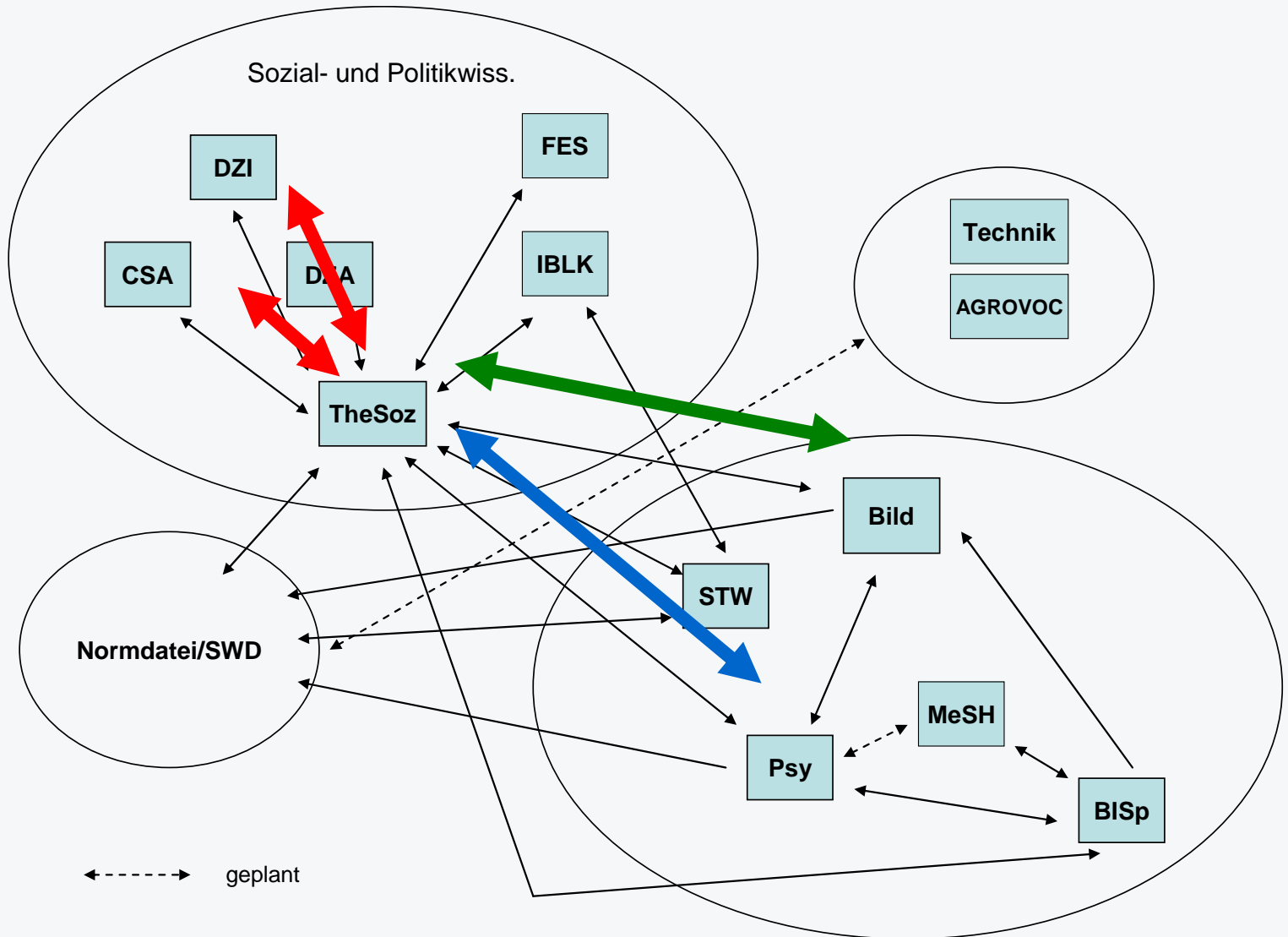
Kompetenznetzwerk Modellbildung und Heterogenitätsbehandlung:

Teilprojekt innerhalb des Kompetenznetzwerks „Neue Dienste,
Standardisierung, Metadaten“ (BMBF)

- Modellbildung: übergreifende Modellbildung für komplexe Informationsinfrastrukturen, u. a. am Beispiel des Wissenschaftsportals „vascoda“ mit allen nachgeschalteten Ebenen
- Heterogenitätsbehandlung: als notwendige Ergänzung zur Standardisierung durch einheitliche Metadaten

Projektlaufzeit September 2004 bis August 2007

Crosskonkordanzen Übersicht



Einsatz von Crosskonkordanzen in infoconnex

Beispiel: Sozialwissenschaften; Metasuche nach „Kulturpolitik“

- Konkordanz TheSoz → DZI:
Kulturpolitik =+ Kultur und Politik
- Konkordanz TheSoz → CSA:
Kulturpolitik =+ Cultural Arts Policy

The screenshot shows the infoconnex search interface. On the left is a navigation menu with categories like 'infoconnex-Suche', 'Wir über uns', 'About us', 'Pädagogik', 'Sozialwissenschaften' (highlighted), 'Psychologie', 'Konditionen / AGB', 'Registrierung', 'Intern', 'Kontakt', and 'Impressum'. The main content area is titled 'Sozialwissenschaften' and shows a search form with the query 'Schlagwort=Kulturpolitik[1738]'. Below the search form, there are filters for 'SOLIS (1738)', 'FORIS (105)', 'DZI SoLit (154)', and 'Nationallizenzen (259)'. The search results are sorted by 'Jahr' (Year) and show 3 results from 2006. Each result includes a checkbox, the year, the title, the database (SOLIS), and availability icons (a crossed-out circle and a 'subito' logo).

	Jahr	Titel	Datenbank	Verfügbarkeit
<input type="checkbox"/> 1.	2006	Managing ethnic diversity in Javakheti : two European models of multilingual tertiary education [...]	SOLIS	
<input type="checkbox"/> 2.	2006	Rekonstruktion kulturpolitischer Handlungslogik mittels der Analyse von Budgetdaten und Pressemitteilungen : eine Idee zur Methodologie soziologischer Politikanalysen [...]	SOLIS	
<input type="checkbox"/> 3.	2006	Der Prager Film in den Jahren der Okkupation 1939-1945 [...]	SOLIS	



Verfügbarkeit der Crosskonkordanzen

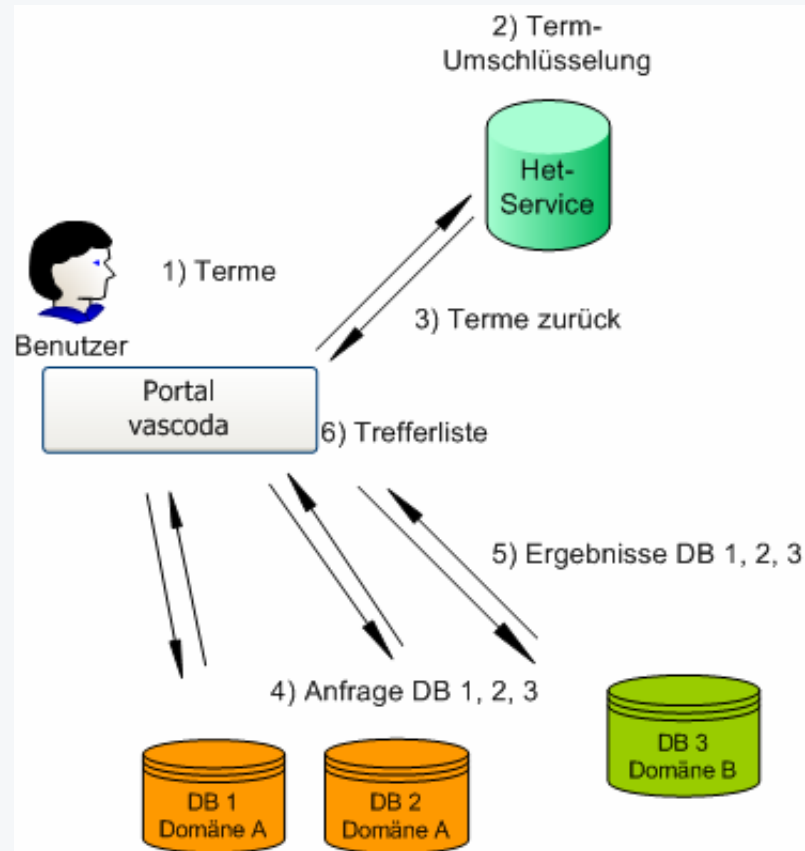
Heterogenitätsservice:

- Rückgabe von Termtransformationen für einen Anfrageterm
- Erste Testimplementierung
- Datenaustauschformat: XML

Weiteres Vorgehen:

- Erweiterung der Datenbasis um die aktuell fertig gestellten Crosskonkordanzen
- Verwendung von Standards für die technische Schnittstelle: SKOS für das Rückgabeformat
- weitere Spezifikation des Funktionsumfangs

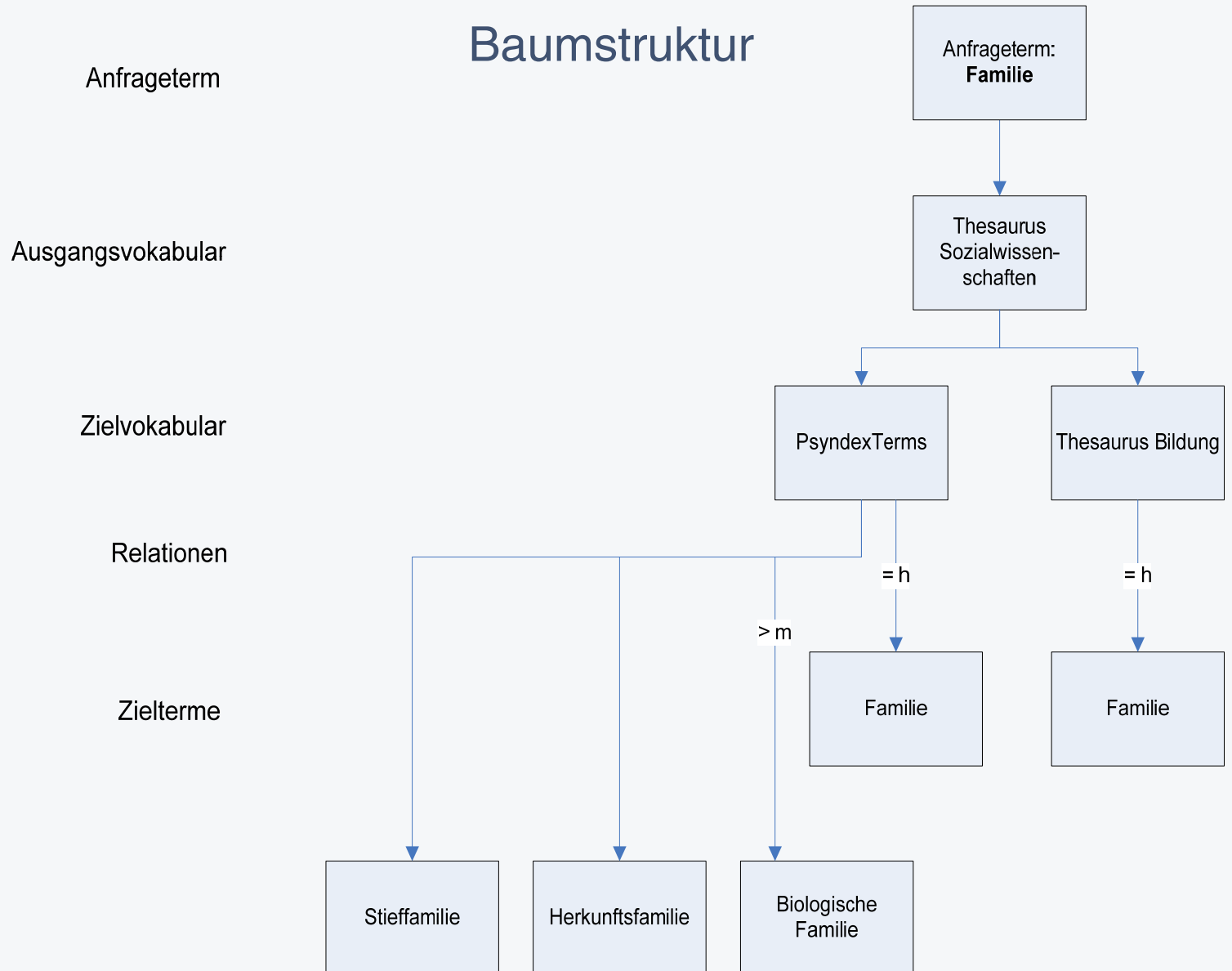
Einsatzszenario (Heterogenitätsservice)



- 1) Nutzer gibt in vascoda Suchterme ein
- 2) Het.-Service schlüsselt übergeb. Terme um
- 3) DB liefert umgeschlüsselte Terme zurück
- 4) Portal setzt Anfrage um und schickt an DB
- 5) Ergebnislieferung pro Datenbank an Portal
- 6) Aufbau der Trefferliste im Portal

Beispiel: Antwort des Heterogenitätsservices

Baumstruktur



- Weitere Crosskonkordanzen
 - Technik
 - Agrovoc
 - Klassifikationen
- Einsatz statistischer Verfahren
 - MeSH-SWD
- Einsatz des Heterogenitätsservice in Sowiport, vascoda
- Anfragebearbeitung an Benutzerschnittstelle
- Qualitative Evaluierung der Termtransformationen



Vielen Dank für Ihre Aufmerksamkeit!

Philipp Mayr, Anne-Kathrin Walter

{walter,mayr}@bonn.iz-soz.de

<http://www.gesis.org/IZ/>