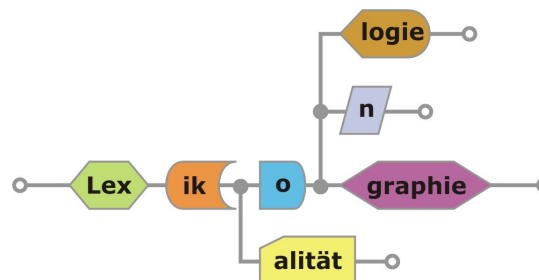


Die Erstellung einer 'Metalemmaliste' zur deutschen Sprache im Rahmen des Projekts

**Wechselwirkungen zwischen linguistischen und
bioinformatischen Verfahren, Methoden und Algorithmen:
Modellierung und Abbildung von Varianz in Sprache und
Genomen**



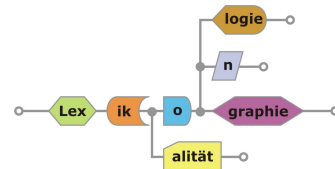
Andrea Rapp, Kompetenzzentrum Trier

www.kompetenzzentrum.uni-trier.de

andrea.rapp@uni-trier.de

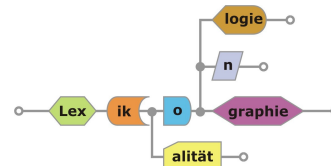
Inhalt

- Vorstellung des Förderprogramms & der Projektpartner
- Vorstellung der Projektstruktur & der Zielsetzung
- transdisziplinäre Forschungsfragen
- Bezugssystem und Teilergebnis für die Sprachwissenschaft: *Metalemmaliste*
- sprachwissenschaftliche Datenbasis: Wörterbuchdaten
- Konzeption und Aufbau der Metalemmaliste
- Nachhaltigkeit durch Kooperation: z.B. TextGrid
- Ausblick: Kollaborativer Ausbau



BMBF-Förderprogramm "Wechselwirkungen zwischen Natur- und Geisteswissenschaften"

- In diesem Förderschwerpunkt arbeiten geistes- und naturwissenschaftliche Fächer in interdisziplinären Forschungsverbünden zusammen.
- In die interdisziplinäre Zusammenarbeit müssen sowohl geisteswissenschaftliche als auch naturwissenschaftlich-technische, mathematische oder informationstechnologische Kompetenz einfließen.
- Sowohl der Einsatz naturwissenschaftlicher Methoden in den Geisteswissenschaften als auch umgekehrt, der Einsatz geisteswissenschaftlicher Methoden in den Naturwissenschaften ist erwünscht, um nicht zuletzt neue Methoden zu entwickeln.





Vorstellung der Verbundpartner

Verbundkoordinator



Kompetenzzentrum
für elektronische Erschließungs-
und Publikationsverfahren in den
Geisteswissenschaften

Prof. Dr. Claudine Moulin, Kompetenzzentrum für elektronische Erschließungs- und Publikationsverfahren in den Geisteswissenschaften, Universität Trier

Verbundpartner



INSTITUT FÜR
DEUTSCHE SPRACHE

Prof. Dr. Ludwig Eichinger, Institut für Deutsche Sprache, Mannheim



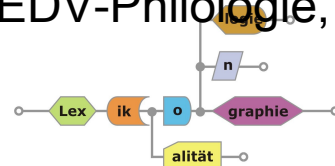
Prof. Dr. Dietmar Seipel, Institut für Informatik, Lehrstuhl für Informatik I, Informationsstrukturen und wissensbasierte Systeme, Universität Würzburg



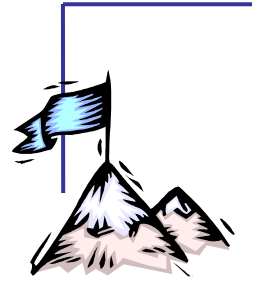
Prof. Dr. Jörg Schultz, Dipl.-Biol., Theodor Boveri-Institut für Biowissenschaften/Bioinformatik, Universität Würzburg



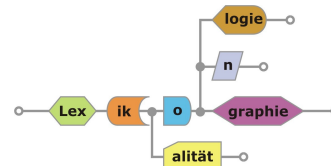
Prof. Dr. Werner Wegstein, Kompetenzzentrum EDV-Philologie, Universität Würzburg



Fragestellung und Ziel

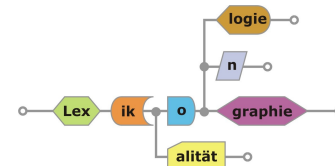


- prinzipielle strukturelle Gemeinsamkeiten
 - zwischen Genomcode und sprachlichem Code
 - biologischer Entwicklung und Sprachentwicklung
- zentrale gemeinsame Kennzeichen (u.a.)
 - Entwicklungsfähigkeit -> Historizität
 - Vielfalt bzw. Varianz der Erscheinungen
- differenzierteres Verständnis der Konzepte und Mechanismen von Entwicklung und Varianz ermöglicht
 - neue und präzisere Verfahren der Informationsgewinnung
 - der Speicherung
 - Bearbeitung und
 - Auswertung der dadurch gewonnenen Daten



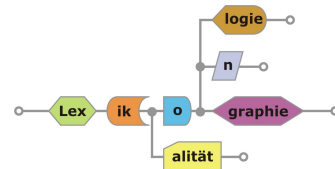
Projektstruktur: 4 Teilvorhaben

- TV1: Basis-Lemmaliste der neuhochdeutschen Standardsprache (Institut für Deutsche Sprache Mannheim)
- TV 2: Klassifizierte Varietäten-Lemmalisten, semasiologische Vernetzung, Erkundungsmodul onomasiologische Vernetzung (Kompetenzzentrum Trier)
- TV3.1: Grammatik der Varianz in Genomen und Sprache, Quantitative vergleichende Analysen (Bioinformatik Würzburg)
- TV 3.2: Technik: Visualisierung, Gridifizierung, Modellierung der Vernetzung (Informatik und Kompetenzzentrum EDV-Philologie Würzburg)
- Projektstart: 1.12.2008



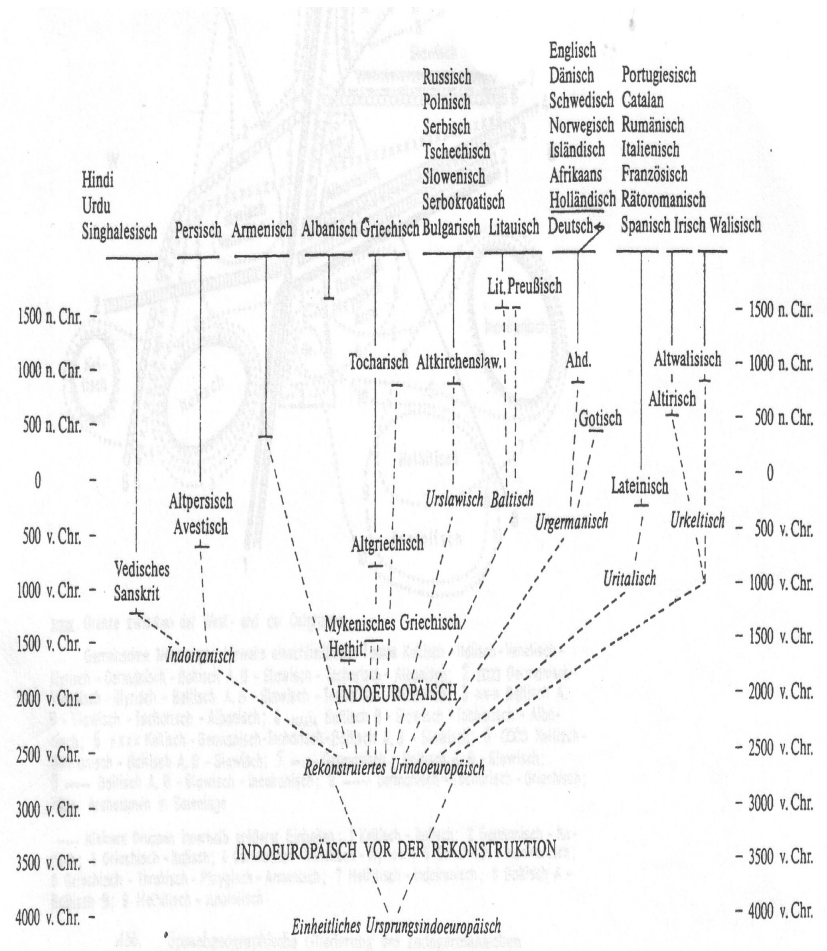
Gemeinsame Forschungsfragen

- Identifikation analoger Mechanismen von Entwicklung und Varianz in Sprache und Genomen
- exemplarisch:
 - Strukturen: Domänen und Morpheme
 - regionale und historische Varianz
 - onomasiologische Varianz
 - Entlehnungsmechanismen

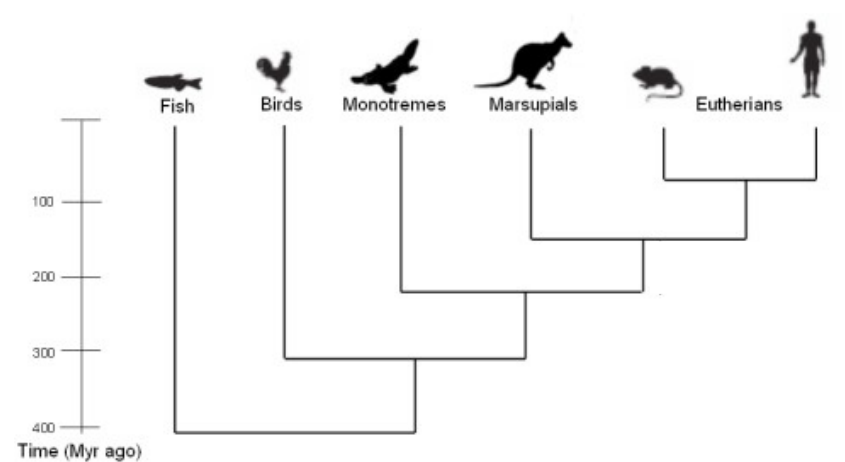


Evolution in der Linguistik und der Biologie

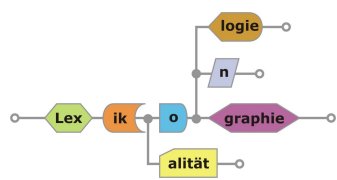
Die Entwicklung der idg. Sprachen



Die Entwicklung der Vertebraten (Wirbeltiere)



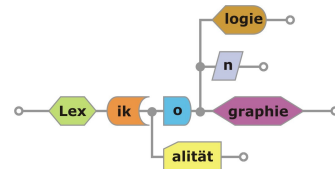
➤ Unterschied der Zeiträume!



Strukturen:

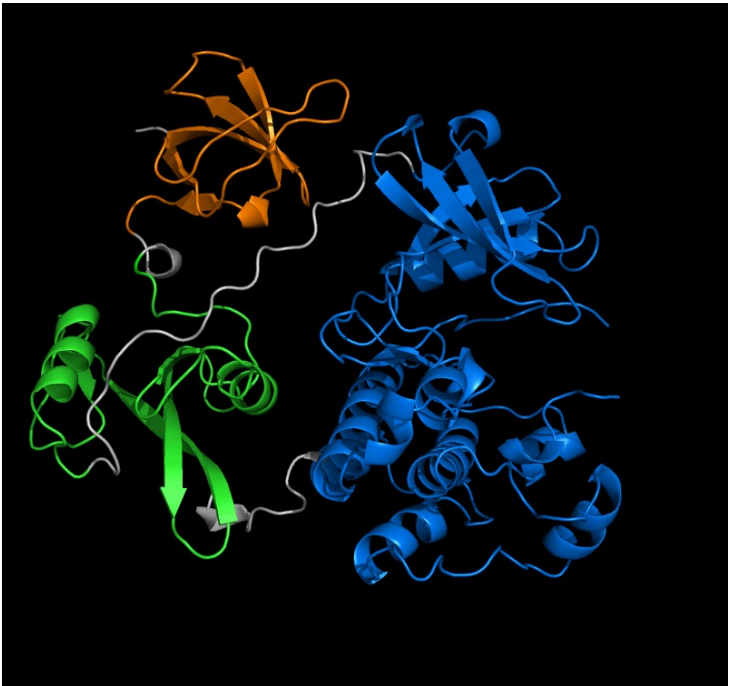
Morpheme

Domänen



Domänen in Proteinen

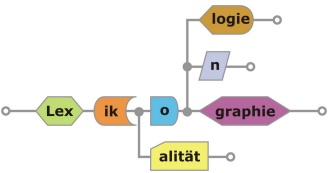
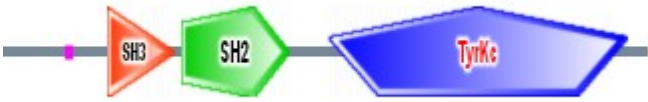
Struktur



Sequenz

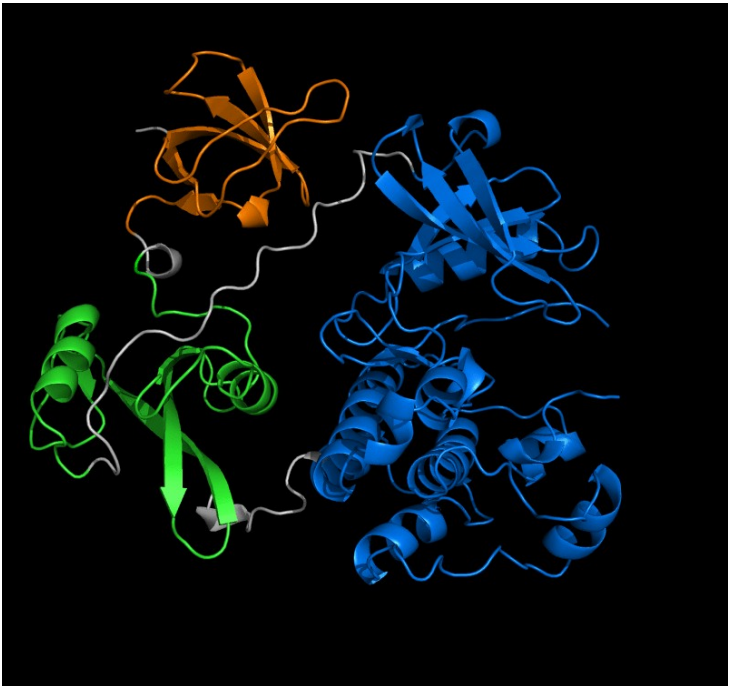
MEPFLLRKRLAFLSFFWNKIWPAGGEPVSDGSWSPDAEFPAEPGSPAPARLFRALSDFRA
RCAGELSVSRGDWLFALQEDGEYLFARRLSSPFCTGLVPITCVAKTSPEPLSDQWPYFGD
ISRTQAQQLLLSAANAPGAFLVRPSESRRGDYSLSVRAQAKVCHYRISMAEDGGLYLQKG
QCFSSLEELLSYYKANWKLLQNPLLQPCMPQKSQEQQDQWERPRSEFVLRRLKGQGGFFGEV
WEGWLWGSVPVAVKVIKSADMKLDDLAQEIRTLKSLRHERLIRLHAVCSTGEPVYIVTEL
MRKGSLLQAYLGPGGRTLSLPLLLSFACQVAEGMGYLECRRIVHRDLAARNVLVGDNLAC
KVADFGLARLLKDDVYSPSSSSKIPVKWTAPEAANYCIFSQKSDVWSFGVFLYEVFTYGGQ
CPYEGMSNHETLQQVMQGYRLPRPPTCPAEVYVLMLECWKGSPEERPAFSVLQEKLGDIS
RCFYP

Domänen



Domänen in Proteinen

Struktur

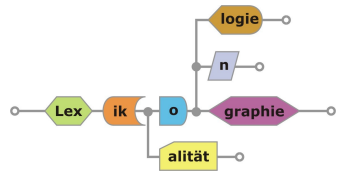
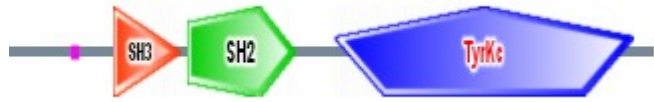


Sequenz

MEPF~~LR~~KRLAFLSFFWNKIWPAGGE~~PVSDG~~SWSPDAEFPAEPGSP~~PAPA~~~~RLFR~~ALSDFRA
RCAGELSVSRGDWLFALQEDGEYLFARRLS~~SP~~FCTGLVPIITCVAKTSPEPLSDQ~~PWY~~FGD
ISRTQAQQLLSAANAPGAFLVRPSESRRGDYSLSVRAQAKVCHYRISMAEDGGLYLQKG
QCFSSLEELLSYYKANWKL~~LQN~~PLLQPCMPQKSQE~~QDQ~~WERPRSEFVLRRKL~~GQ~~GGFFGEV
WEGWL~~GS~~VPVAVKVIKSADMKLDDLAQEIRTLKSLRHERLIRLHAVCSTGEPVYIVTEL
MRK~~GS~~LQAYLG~~GP~~GGRTLSLPLLLSFACQVAEGMGYLECRRIVHRDLAARNVLVGDNLAC
KVADFG~~LAR~~LLKDDVYSPSSSSKIPVKWTAPEAA~~NYC~~IFSQKSDVWSFGVFLYE~~VFT~~YTGQ
CPYEGMSNHETLQQVMQGYRLPRPPTCPAEVYVLMLECWKGSPEERP~~AF~~SVLQEKLGDIS
RCFYP

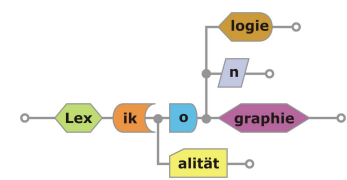
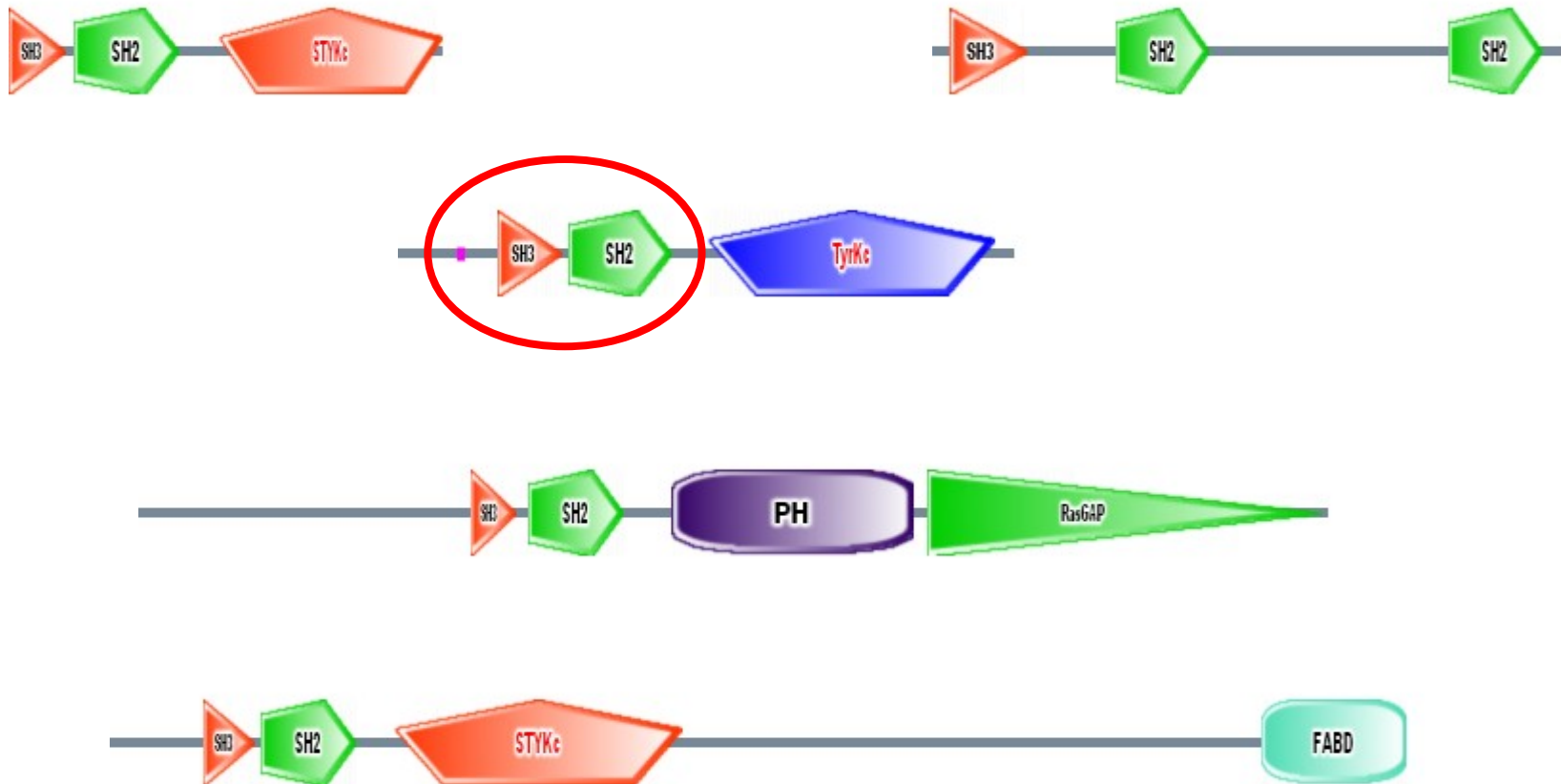
Hidden Markov Modelle

Domänen

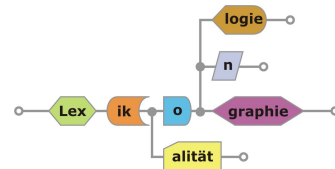
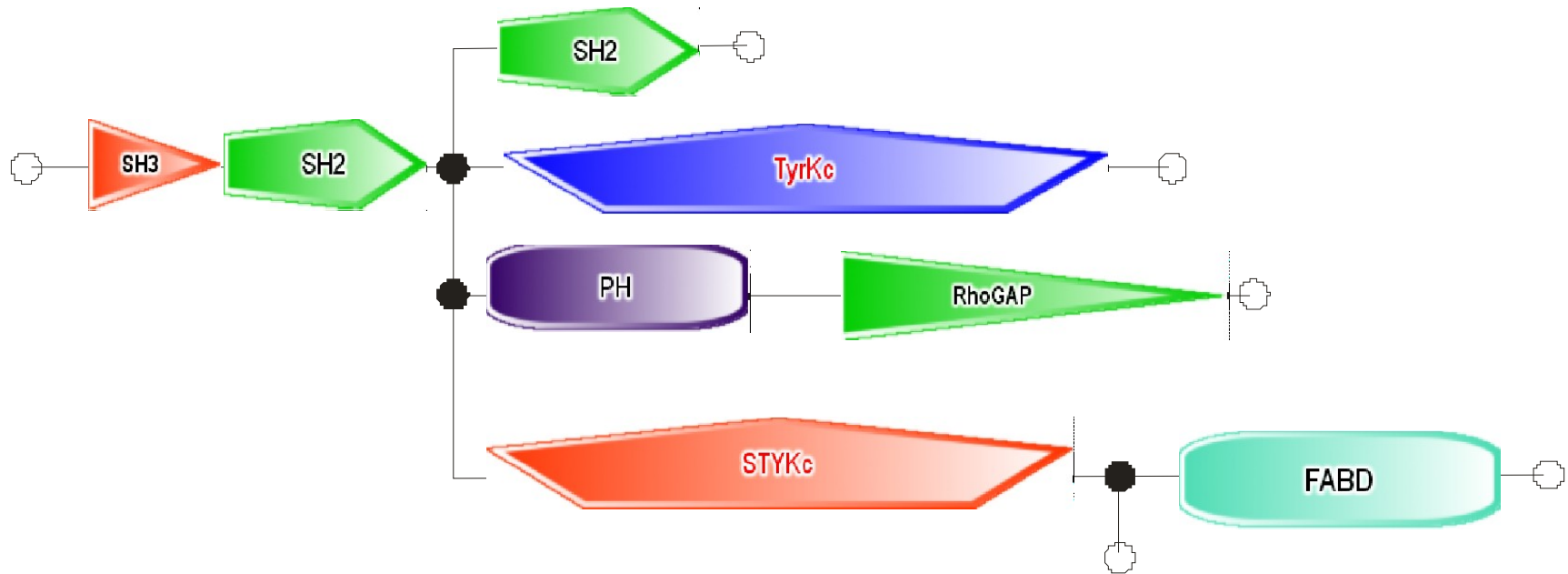


Lexikon von Proteindomänen

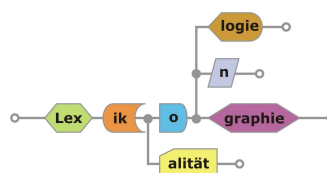
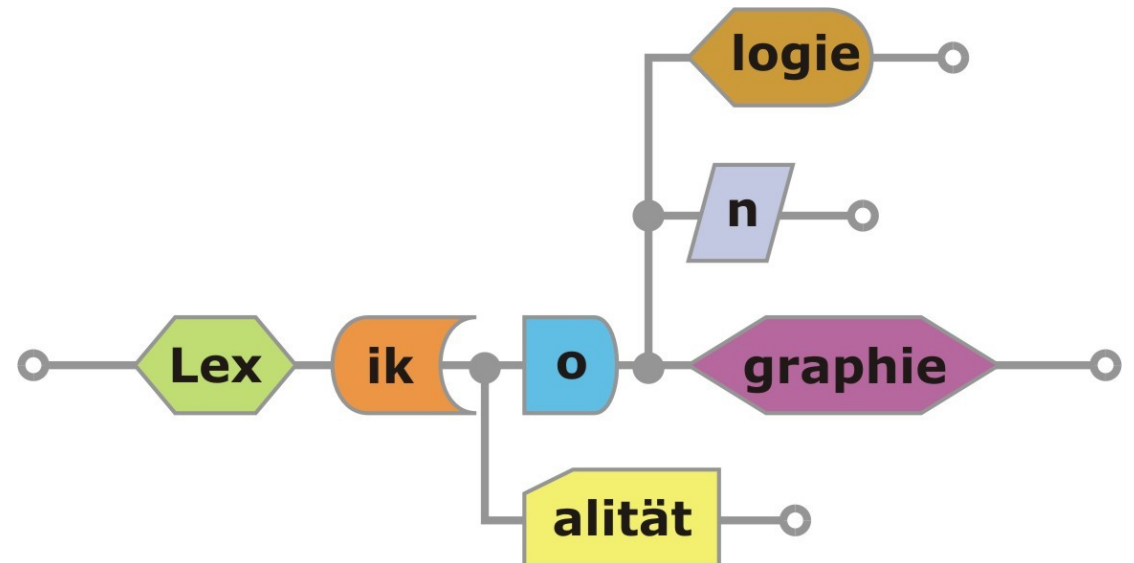
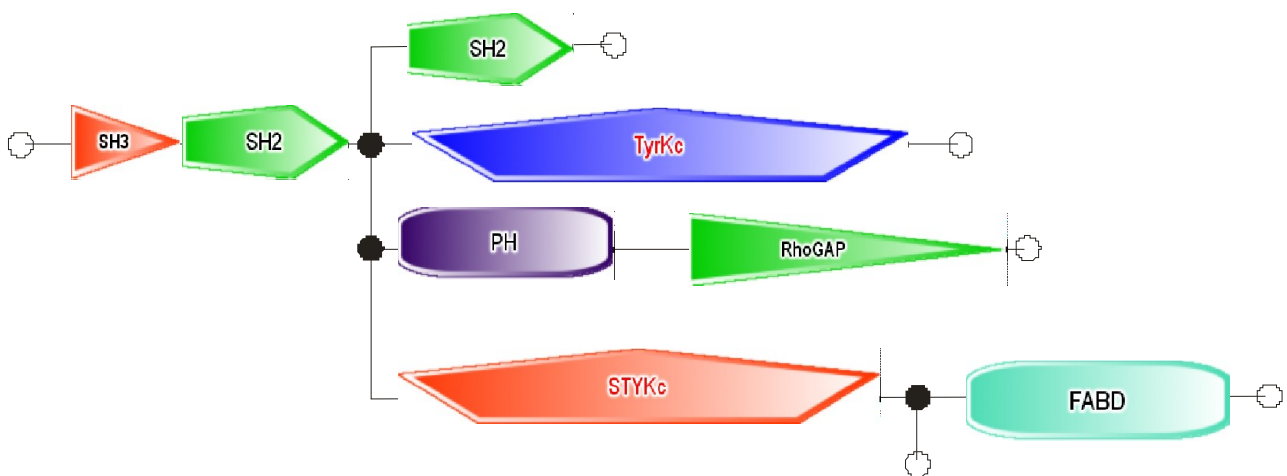
Weitere Einträge mit Proteinstamm?



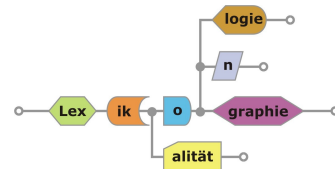
Domänen in Proteinen



Domänen und Morpheme



regionale und historische Varianz



Varianz in der Sprache am Beispiel ‚Brombeere‘



Parameter: semasiologisch (ausdrucksseitig)

idg.	<i>*bher-</i>	hervorstehen, etwas Spitzes (z.B.
idg.	<i>*bha-</i>	glänzen Dornen)
ahd.	<i>brāma, prāma</i>	Dornbusch
mhd.	<i>brāme</i>	Dornbusch
ahd.	<i>brām-beri</i>	Brombeere
mhd.	<i>brām-ber</i>	Brombeere
nhd.	<i>Brom-beere</i>	

Brombes

Kleve

Brommelte

Düsseldorf

Bromele

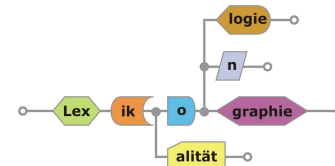
Bonn

Brambel

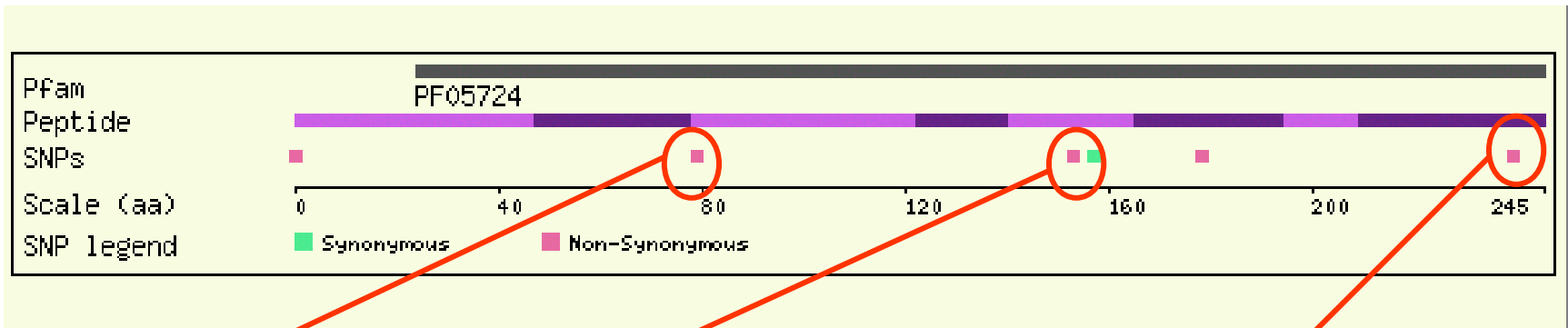
Saarlouis

Bromere

Schlettstadt



Thiopurine S-methyltransferase loss-of-function Mutationen

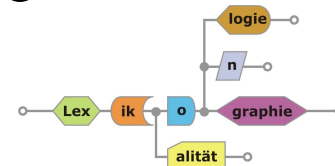


TMPT*2 A->P

TMPT*3A A->T
Häufigste Variante in
Amerikanisch / Kaukasisch

TMPT*3C Y->C
Häufigste Variante in
Afro-Amerikanern

Regional verteilte Varianten beeinflussen Dosierung
von Medikamenten



Varianz in der Sprache am Beispiel ‚Brombeere‘



Parameter: onomasiologisch (inhaltsseitig, semantisch)

nhd. *Brom-beere*

Dornbeer

Kratzbeer

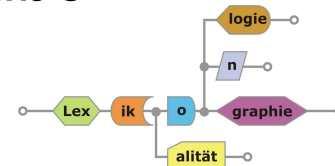
Schmärzbeer, Schmäsbeer, Schmelzbeer

Heckenbeer, Heckbes, Heckebraimcher

Tranbeer, Bachbeer,

Schwa(r)zbeer, Schmoazbel, Schwosbat, Woschbert

Perdsbeer, Fuchsbeer, Katzentapen, Katzedoube



Neue Einheit – gleiche Funktion

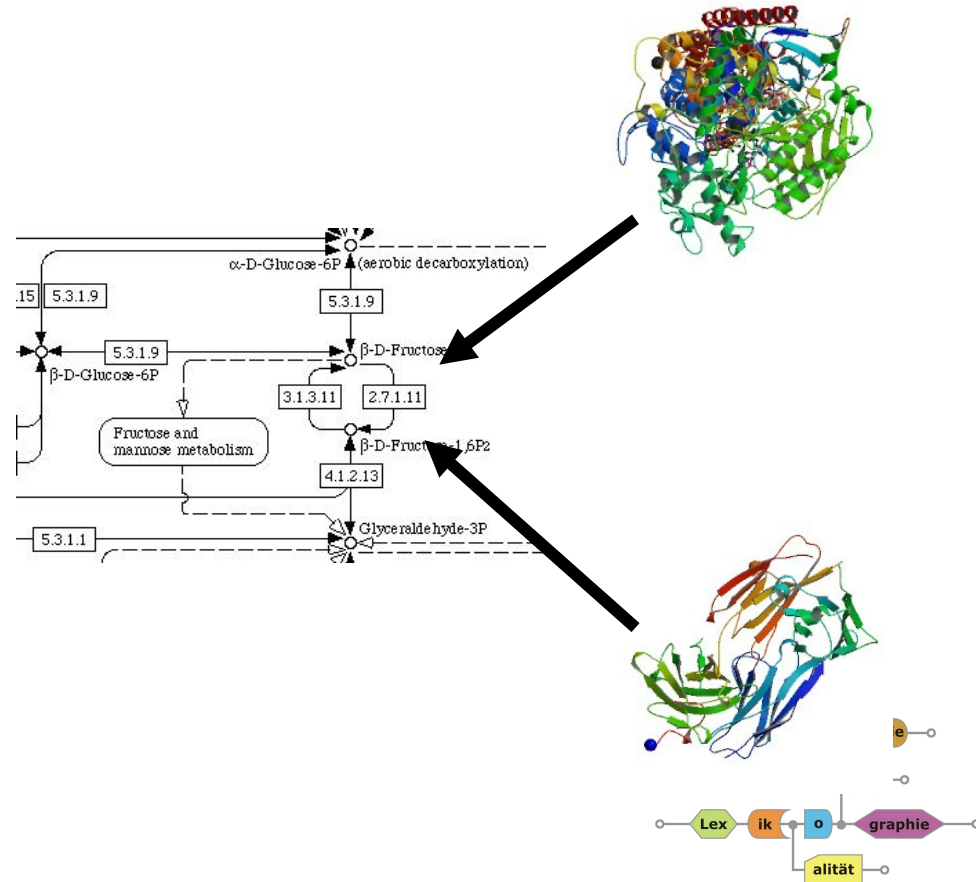
Onomasiologische
Varianz

Nicht orthologe Gen-
Ersetzung

Brombeere

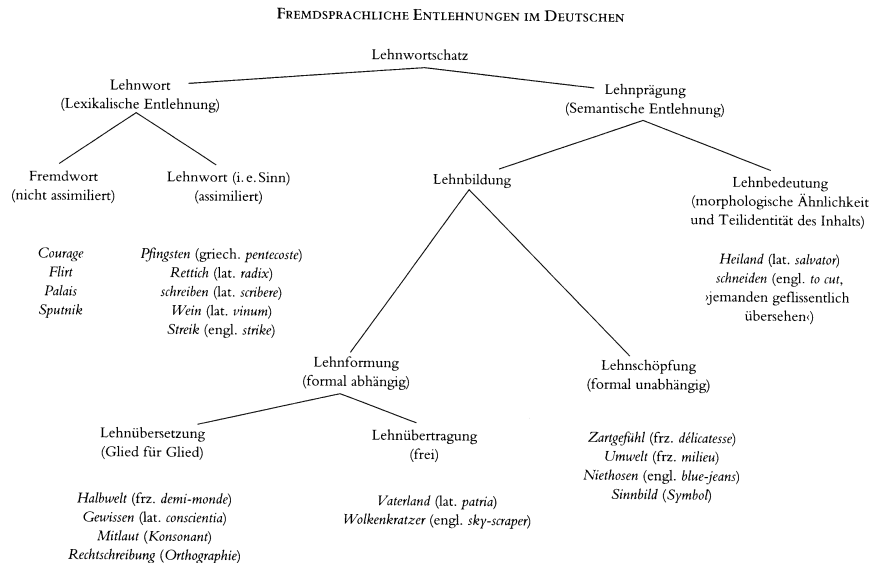


Fuchsbeere

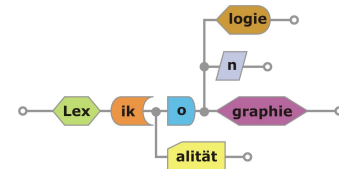
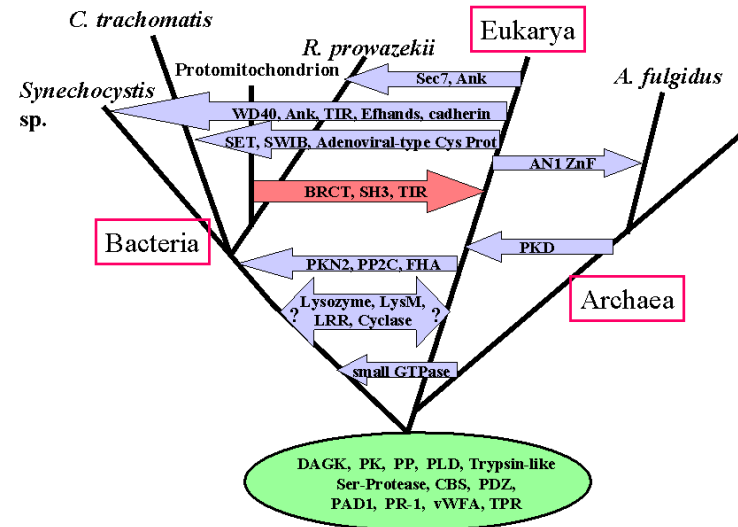


Horizontaler Transfer vs. Entlehnung

Entlehnungsmechanismen



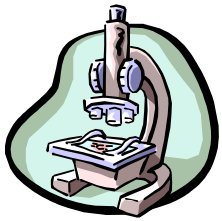
Horizontaler Gentransfer



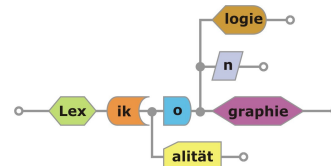
Empirisch-quantitative Evaluierung



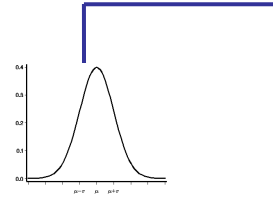
Gehorchen die Prozesse
gleichen empirischen Gesetzen?



Benötigt werden:
o Infrastruktur
o Daten



Datengrundlage



➤ Biologie

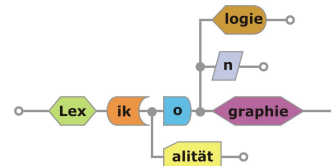
- Verbundpartner beteiligt an der Erarbeitung der benötigten Datenbasis
- Genomprojekte (Mensch, Maus)

→ Bezugssystem vorhanden

➤ Sprachwissenschaft

- Verbundpartner beteiligt an der Erarbeitung der benötigten Datenbasis
- digitale Wörterbücher (Lemmalisten)
- Dialektdatenbanken

→ Bezugssystem fehlt: Meta-Lemmaliste



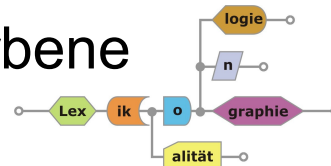
standardisiertes Bezugssystem:



Metalemmaliste

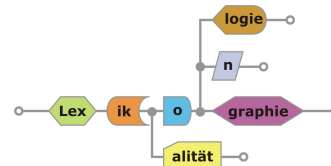
Metalemma (nhd. Standard, öffentlicher Sprachgebrauch)

- gemeinsames Drittes
- verbindet unterschiedliche Wortparadigmen der Varietäten
- Schnittstelle für externe Informationssysteme (z.B. Ontologien)
- Kernbestand vorhanden
 - Auf der Basis des IDS-Referenzkorpus Liste der 30.000 häufigste Wörter als Grundlage für die Meta-Lemmaliste
 - Teil-Lemmalisten (als Teil-Bezugssystem) zu spezifischen Bereichen, z.B. Sprachstufe Mittelhochdeutsch oder westmitteldeutsche Dialekte
- Varietätenlemmata müssen mit Metalemmata in Bezug gesetzt werden
- Bildung von Pseudo-Metalemmata, z.B. für ausgestorbene Wörter



Bezugssystem: DeReWo-Liste

- ermittelt auf der Basis des **Deutschen Referenz-Korpus** am Institut für Deutsche Sprache (IDS) Mannheim
- Korpus mit 3,2 Milliarden Textwörtern
- Texte ab 1956
- belletristische, wissenschaftliche und populärwissenschaftliche Texte, Zeitungstexte, weitere Textarten
- laufend erweitert
- urheberrechtlich abgesichert
- Kern (pragmatisch gewählt): die 30.000 häufigsten Wörter – "beliebige" Erweiterung möglich
- Ausbau, Modifikation



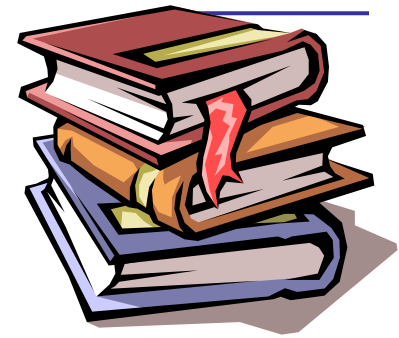


Auszug aus der
DeReWo-Liste
mit einfacher
XML-Struktur:
Zähler, Lemma,
Häufigkeitsgruppe

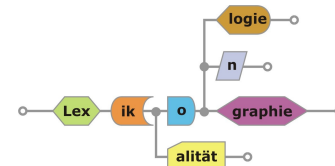
```
V:\glmetall\DATEN\DeReWo\derewo.xml
<entry id="29970" lemma="Karriereleiter" group="17" />
<entry id="29971" lemma="detoniert" group="17" />
<entry id="29972" lemma="Passagierzahl" group="17" />
<entry id="29973" lemma="Götterdämmerung" group="17" />
<entry id="29974" lemma="Exzellenz" group="17" />
<entry id="29975" lemma="Elegie" group="17" />
<entry id="29976" lemma="hervorholen" group="17" />
<entry id="29977" lemma="Wasserglas" group="17" />
<entry id="29978" lemma="Ursachenforschung" group="17" />
<entry id="29979" lemma="Sozialgeschichte" group="17" />
<entry id="29980" lemma="Reinigungsmittel" group="17" />
<entry id="29981" lemma="Mitbürgerin" group="17" />
<entry id="29982" lemma="Fontäne" group="17" />
<entry id="29983" lemma="Dokortitel" group="17" />
<entry id="29984" lemma="Zentralorgan" group="17" />
<entry id="29985" lemma="Vereinskasse" group="17" />
<entry id="29986" lemma="Okkupation" group="17" />
<entry id="29987" lemma="Monopoly" group="17" />
<entry id="29988" lemma="Leerstelle" group="17" />
<entry id="29989" lemma="Konfirmand" group="17" />
<entry id="29990" lemma="Kerzenlicht" group="17" />
<entry id="29991" lemma="Filmreihe" group="17" />
<entry id="29992" lemma="Einsender" group="17" />
<entry id="29993" lemma="Bastille" group="17" />
<entry id="29994" lemma="Badesaison" group="17" />
<entry id="29995" lemma="ausspionieren" group="17" />
<entry id="29996" lemma="appetitlich" group="17" />
<entry id="29997" lemma="allerwenigste" group="17" />
<entry id="29998" lemma="Verkehrsstrom" group="17" />
<entry id="29999" lemma="Philologe" group="17" />
</derewo>
```

Fertig

Varietätenlemmalisten aus dem Wörterbuchnetz



- Deutsches Wörterbuch von Jacob Grimm und Wilhelm Grimm
- Mittelhochdeutsches Wörterbuch von Benecke Müller und Zarncke
- Mittelhochdeutsches Handwörterbuch von Matthias Lexer
- Findebuch zum mittelhochdeutschen Wortschatz
- Pfälzisches Wörterbuch
- Rheinisches Wörterbuch
- Wörterbuch der elsässischen Mundarten
- Wörterbuch der deutsch-lothringischen Mundarten
- Goethe-Wörterbuch
- Lexicon der Luxemburger Umgangssprache von J. F. Gangler
- Wörterbuch der luxemburgischen Mundart






Bezugssysteme: Wörterbücher



mhd. Wbb
(BMZ und
Lexen)

brâmbere   stn. *brombeere*. *sumerl.* 40,70. 56,77. 57,53. 1, 258. *daʒ hulfe in niht ein brâmbere* *Mone altd. schausp.* 3,44

brâm-ber  stn. (→ I. 104^a) *brombeere* *MONE schausp.* EILH. *swarz gevertet als ein zitic brâmbere* *TROJ.* 32743; *bromber* *HPT.*

hist.
nhd. Wb.
(Grimm)

BROMBEERE, *f. rubum*, *brambeere*, *schwankt in branbeere*,

Bd. 2,

branbire, *bramber*, *braunbeere*, *brobeere*, *braubeer*, *brommer*. *DASY.*
brombeer morum rubi.

Pfälz. Wb

Brombeere *f.*: wie schd., zur Verbr. vgl. K. 68, s. auch Karte :
in DWA X, *Brambee* (*brambē*) [verbr. im pfälz.-elsäss. Grenz-]

Rhein. Wb

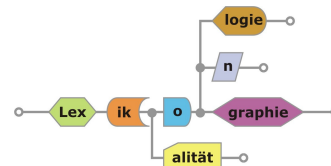
Bram-beere Formen. mit erhaltenem *-beere* u. *brâm-* > *brom-*:
brombēs, Pl. *-z* Klevld [brom□lber□n Rees-Brünen, brøm□lbēs :

Els. Wb

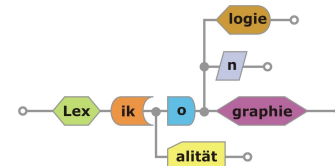
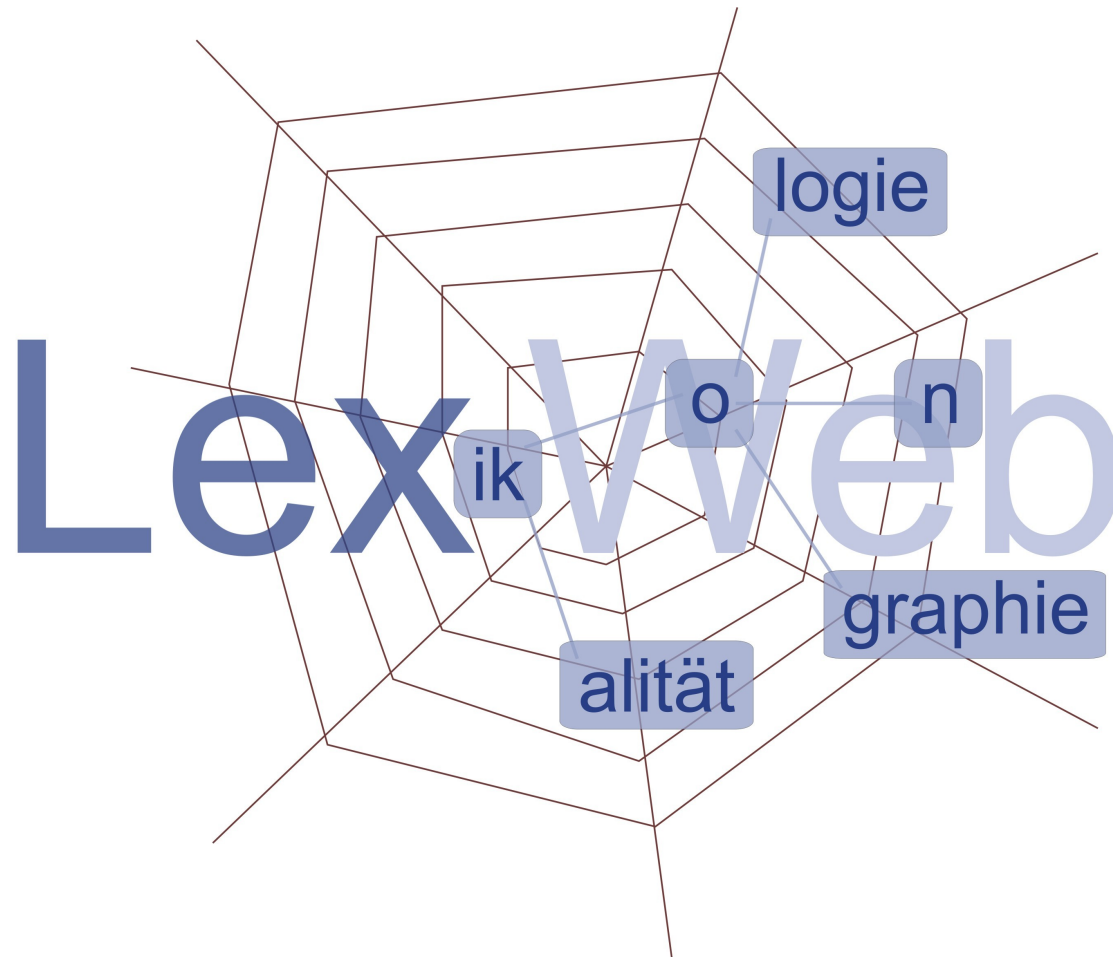
Brambeer, *Brombeer*, *Brombere*, *Bromer(e)*, *Bromert*, *Brämer*,
Bromter, *Blombeer*, *Blomber*, *Blomere* [Prämpêr Ndröd. *Wingen Lohr*;
Prompter Co.; Prümpêr *Ensisl. Urbis Su. Mutzig Str. K. Z. Betschd.*;
Prümpêr *Pfetterhsn.*; Prômêr *Roppenzw. Hürzf.*; Prômêr *Co. Rchw.*

Lothr. Wb

Brämer [brémær, bréimær *Si.*] *m. Brombeerstrauch.* – *els. 2, 189*
Brämer; mhd. *brâmbere*.




Aufgabe: Herstellung philologisch abgesicherter Vernetzung (semasiologisch)




Mittelhochdeutscher Wörterbuchverbund

Mittelhochdeutsches Handwörterbuch von Lexer



Vernetzungsinformation

brâm-ber  *stn.* (→ I. 104^a) *brombeere* *MONE schausp.* EILH. 1717. *swarz*
geverwet als ein zitic brâmbler TROJ. 32743; bromber HPT. 5,14;

brâm-ber  *stf.* *brombeerstrauch* MGB. 330,25.

brâm-ber, <i>stn.</i>			
Buch	brâmbler, <i>stn.</i>	FiB	
Symm.	brâmbler, <i>stn.</i>	BMZ	brom-ber, <i>stn.</i> LEX
brâm-ber, <i>stf.</i>			
Buch	brâmbler	MLEX	

Mittelhochdeutsches Wörterbuch von Benecke/Müller/Zarncke

brâmbler   *stn.* *brombeere.* *sumerl.* 40,70. 56,77. 57,53. *Schmeller* I,
258. *da3 hulfe in niht ein brâmbere* *Mone altd. schausp.* 3,446.

brâmbler, <i>stn.</i>			
Buch	brâmbler, <i>stn.</i>	FiB	brâm-ber, <i>stn.</i> LEX
Trans.	brom-ber, <i>stn.</i>	LEX	

Trierer Findebuch

brâmbler   *stn.* ENIK. HVNST.

brâmbler, <i>stn.</i>			
Buch	brâmbler, <i>stn.</i>	BMZ	brâm-ber, <i>stn.</i> LEX
Trans.	brom-ber, <i>stn.</i>	LEX	

Brombeere f.: wie schd., zur Verbr. vgl. K. 68, s. auch Karte Brombeere in D

X, *Brambee* (*brambē*) [verbr. im pfälz.-elsäss. Grenzgebiet PS-Lembg KL-Frankst NW-Weidth Hardbg Dürkh Wachh LU-Alsh/Gr Limbghf LA-Altd GH-Hayna Rh'zab], *Brombee* (*brom-*) [KU-SchönbG HB-Brenschb PS-Geisbg KL-Hütschhs Nd'mohr O'sulzb Rockhs schmales Gebiet zwischen Kaislt u. Frankth FR-Albsh Flomh Eppst Raum zwischen Spey u. Germh (neben *Brambee*) LA-Roschb Mörzh Bergz verbr. Gal Buch], *Brombie* [PS-O'simt Gal-Sap Buch-A'fratautz], *Brumbee* (*brumbē*) [IB-Aßw PS-Schöna Busbg Schindhd Erfw Geisbg FR-Heßh NW-Erph Haardt Neustdt LA-Frankw BZ-Ingh Klingmst Germh GH-Hagb Don Franztal Mercydorf Gal-Brig Hohenb Stanin], *Brembee* (*brēmbē*) [verbr. im Raum zwischen St. Wendel u. Kaislt sowie in der gesamten Südpfalz KU-Diedk Horschb RO-Schneebghf Münchw LA-Wey Edk Don-Lenauh Gert Hatzf Gottlob Werb Gal-Moosbg Bruckenth Ottenhs Rosenbg Fehlb], *Brembie* [Kus KU-Etschbg Theisbgstg HB-Alth IB-Bebh Pirmas PS-Erlbn Don-Alexanderhs Tschest St. Andreas Buch-Tereblestie], *Brimbee* [PS-Merzalb BZ-Annw Rambg], *Blambee* [verbr. im Raum südl. Lu'haf (vgl. [BERTRAM](#) 141)], *Blombee* [GH-Sondhs], *Blembee* [GH-Knitth Bellh], *Brame* »Zweibrücker Ggd., Westrich vielfach« ([WILDE](#) 34) LU-Friesh], *Breme* (*brēm□*), selten *brēm□*) [verbr. nördl. WPf NPf nördl. VPf BZ-Billh verbr. Don Gal Buch], *Breme't* (*brēm□d*) [KL-Mölschb W'lein], *Bremde* (*brēmd□*) [im Gebiet nordwestl. KU-Wolfst], *Brime* (*brīm□*) [RO-Hallgt O'mosch Nd'mosch Als], *Brumel* (*brum□l*) [PS-Rodalb BZ-Apphf Kling Mühlhf Germh GH-Hagb], *Bremel* (*brēm□l*) [KB-Stauf FR-Carlsbg A'lein NW-Ungst in einem Gebiet nordwestl. FR-Grünstdt Don-Ulmbach Liebling Gal-Zboiska], *Breme'l* [KB-Rams], *Blame* (*blām□*) [FR-Flomh NW-Hardbg Wachh], *Bleme* (*blēm□*) [im Gebiet um FR-Grünstdt (zusammen mit *Blame*)], *Flambee* [LU-Schauh Assh Iggh NW-Forst Deidh Ruppbg LA-Queichh], *Flame* [im Raum Frankth (zusammen mit *Brombee*)], *Flome* [FR-Bockh LU-Opp], *Flembee* [GH-Zeisk Hördt], *BrÄfl* [GH-Nd'lustdt O'lustdt] (dasselbe Wort auch in dem etwa auf gleicher Höhe östlich des Rheins gelegenen Rußheim). Zs. Heckenbrombeere. Syn.: Brummels-, Dorn-, Heckenbeere, Heckenpraume, Himbeere, Katzentapen, Moz, Schwarzbeere. RA.: 's is so billig wie *Bremere* [KU-Schmittw/O, verbr.]. *Bald is es gut, mei Kind -- ganz bald, dann giehmer* (gehen)

Brombeere

Buch	Brämer LWB	<i>Beer(e)</i> EWB	Bram-beere RhWB
------	----------------------------	------------------------------------	---------------------------------

Brombeeren-kirbe

Buch	Brambeerkirb RhWB
------	-----------------------------------

Brombeeren-saft

Buch	<i>Saft</i> EWB
------	---------------------------------

Brombeeren-wein

Buch	Brambeerwein RhWB
------	-----------------------------------

Promenade

Buch	Promenade RhWB
------	--------------------------------

promenieren

Buch	promenieren RhWB
------	----------------------------------

pro nobis

Buch	pronobissen RhWB
------	----------------------------------

Pröpeler

Buch	brebeln RhWB
------	------------------------------

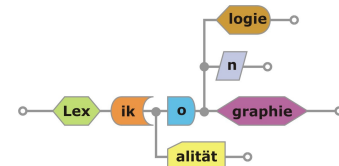
Propeller

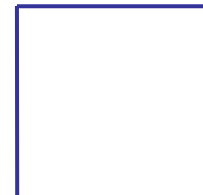
Buch	Propeller RhWB
------	--------------------------------

pröpel'n

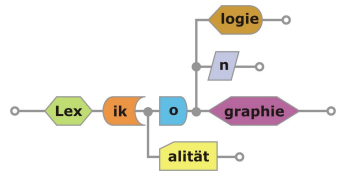
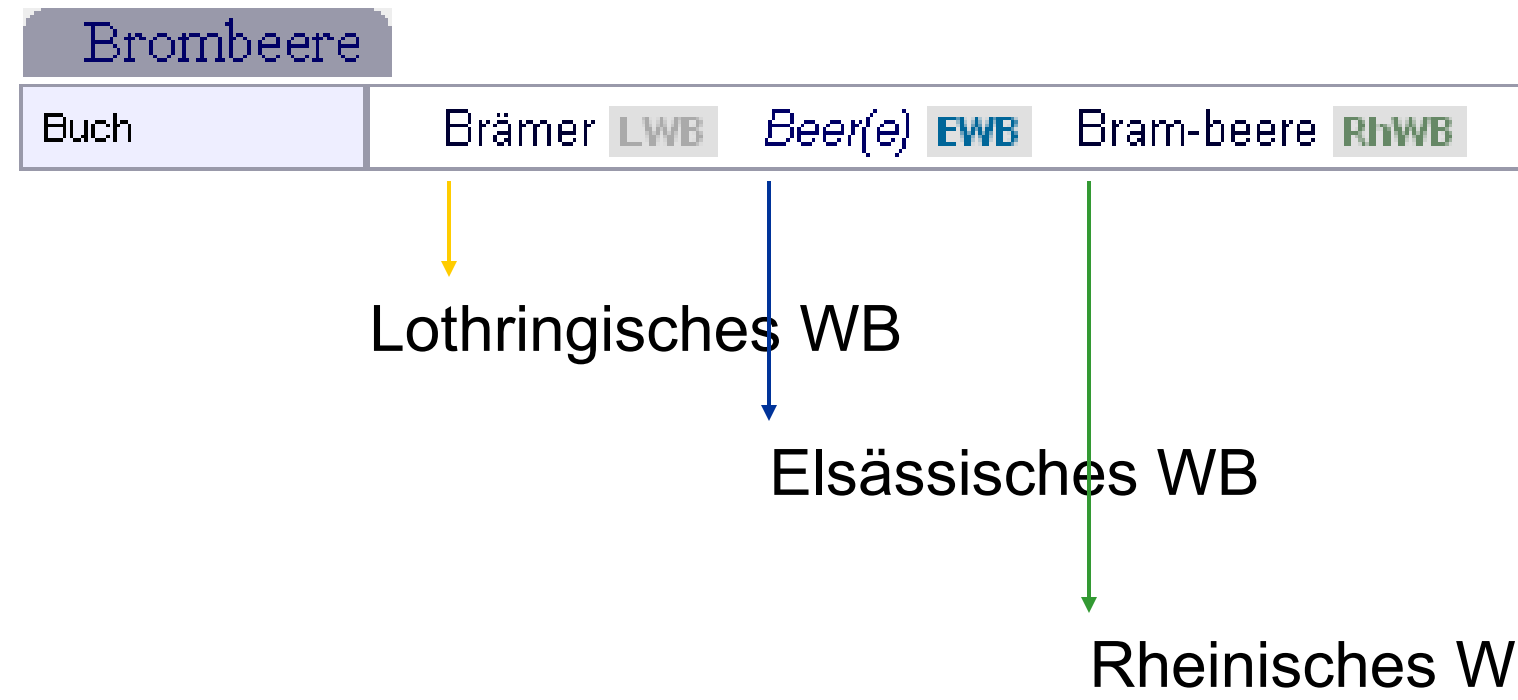
Buch	<i>Prob(e)</i> EWB	PROPELN, v. DWB	brebeln RhWB
------	------------------------------------	---------------------------------	------------------------------

Brombeere im Pfälzischen Wörterbuch mit Vernetzungsinformationen

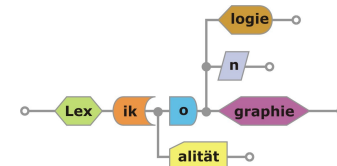
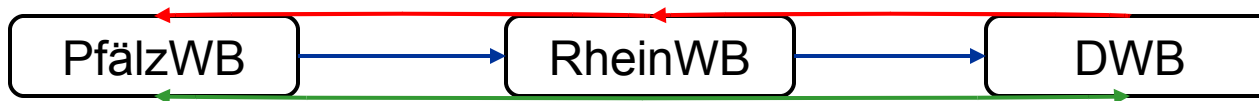




Verweise aus dem Pfälzischen in andere Wörterbücher, übersichtlich aufgelistet und kategorisiert im Vernetzungsfenster



Automatische Verfahren zur Verdichtung der Netzstruktur
z.B. Ausbau **vorhandener** Querverweise durch **Symmetrie** und
transitiven Abschluss

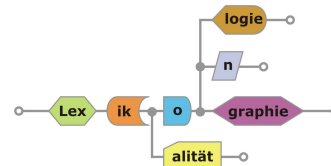


Brämer [brémar, bréimer *St.*] *m. Brombeerstrauch.* – els. 2, 189 Brämer; m. w. brämber.

Lothringische Wörterbuch, Stichwort ‚Brämer‘, dort expliziter Verweis ins Elsässische

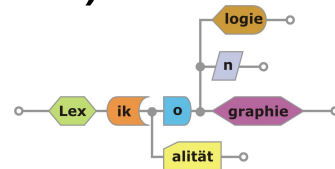
Brämer	
Buch	Bram EWB
Symm.	Brombeere PFWB

Verweis ins Elsässische, vom Computer erzeugter „Rück“-Verweis (symmetrischer Verweis) ins Pfälzische



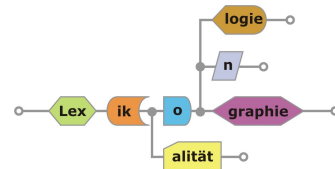
Verbesserung der Zugriffsmöglichkeiten auf Vernetzungsinformationen

- für Rechercheanfragen: zentraler Einstieg über Standard = Metalemma als zentraler Knotenpunkt
- kategorisierte Informationen über Varietäten
- Kategorien z.B.:
 - Zeit
 - Raum
 - Quelle
 - Qualität des Verweises (geprüft/ungeprüft)



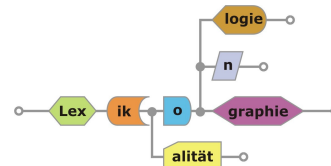
Anwendungs- und Einsatzmöglichkeiten für linguistische Seite des Projekts

- neue und präzisere Verfahren der Informationsgewinnung
- Varianz kann z.B. für phonologisch-graphonematische, morphologische und lexikalische Analysen des Wortschatzes über eine Meta-Lemmaliste ausgeglichen werden, ohne dass dadurch das sprachliche Basismaterial verfälscht wird
- Recherche über Metalemmaliste ermöglicht Suche auf varianten Korpora über standardisierten Zugang
- Verbesserung von Rechercheergebnissen
- Lemmatisierung varianter Texte
- ...



Nachhaltigkeit

- z.B. durch Kooperation mit dem BMBF-geförderten Projekt TextGrid: Community-Plattform und Virtuelle Forschungsumgebung in den Geisteswissenschaften
- freier Zugang zu Forschungsdaten und Werkzeugen
- Nachhaltigkeitskonzepte durch
 - Open Access, Open Source
 - Community-Einbindung, Networking
 - Etablierung von Geschäftsmodellen
- z.B. durch Etablierung der Metalemmaliste als Virtuelle Forschungsumgebung in der Community



DAS WÖRTERBUCH-NETZ

© 2006/2007 ein Projekt des Kompetenzzentrums für elektronische Erschließungs- und Publikationsverfahren in den Geisteswissenschaften an der Universität Trier


 Wörterbuchnetz-
Homepage

Stichwortsuche im gesamten Wörterbuchnetz

Suche starten

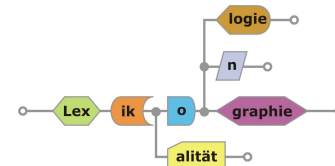
Web-Service
entwickelt
durch



Die Stichwortsuche nach br?mbeere ergab 12 Treffer

Bram-beere	RheinWB	PfälzWB	brambeere	DWB	Brem-beere	PfälzWB	Brim-beere,	RheinWB		
Brimbeere	PfälzWB		Brombeere	PfälzWB	GWB	Brombeere,	RheinWB	brombeere	DWB	DWB
Brum-beere	PfälzWB									

TextGrid Webservice:
Parallele Suche über alle Wörterbücher





Einbindung in TextGridLab

TextGridLab

File Edit XML Search Tools Window Help

Woerterbuchnetz-Suche Project / User Administration Research

Dictionary Search Mask

Keyword: Br?mbeer?

Limit: 50

Searchtype: ☐ exact ☒ fuzzy

☐ linked lemmas only

Dictionary Search Results View

Searchlemma: Br?mbeer?

12 results

Lemma/Dictionary	First Line
Bram-beere	...
RhWB	...
PfWB	...
brambeere	...
DWB	brambeere f. rubum, s. bromt
Brem-beere	...
PfWB	...
Brim-beere,	...
RhWB	...
Brimbeere	...
PfWB	...
Brombeere	...
PfWB	...
GWB	...
Brombeere,	...
RhWB	...
brombeere	...
DWB	brombeere f. rubum, brombee
DWB	brombeere f. rubus, brombee
Brum-beere	...
PfWB	...

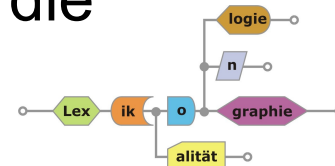
Dictionary List:

- ☒ Dictionaries
 - ☒ Sprachstadien übergreifend
 - ☒ Deutsches Wörterbuch von Jacob und Wilhelm Grimm
 - ☒ Autoren Wörterbücher
 - ☒ Goethe-Wörterbuch
 - ☒ Dialekt Wörterbücher
 - ☒ Elsässisches Wörterbuch
 - ☒ Lothringisches Wörterbuch
 - ☒ Pfälzisches Wörterbuch
 - ☒ Rheinisches Wörterbuch
 - ☒ Rheinisches Wörterbuch Nachträge
 - ☒ Mittelhochdeutsche Wörterbücher
 - ☒ Mittelhochdeutsches Handwörterbuch von M. Lexer
 - ☒ Mittelhochdeutsches Wörterbuch von G.F. Benecke, W. Mueller und F. Zarncke
 - ☒ Findebuch zum mittelhochdeutschen Wortschatz
 - ☒ Nachträge zum Mittelhochdeutschen Handwörterbuch von M. Lexer
 - ☒ Luxemburger Wörterbücher
 - ☒ Wörterbuch der Luxemburger Umgangssprache
 - ☒ Wörterbuch der Luxemburger Mundart

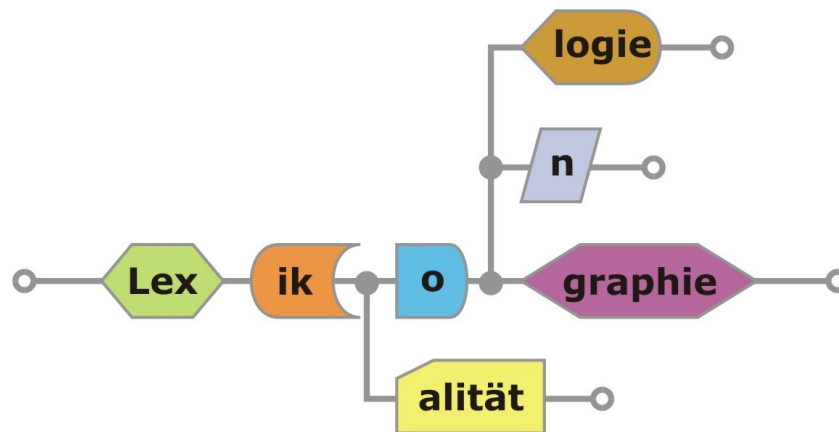
Ausblick



- Metalemmaliste in einer Virtuellen Forschungsumgebung, z.B. für Wörterbuchkanzleien
- Framework zur effizienten Recherche in allen angeschlossenen digitalen Wörterbüchern, z.B. bei der eigenen Artikelarbeit
- Möglichkeit zur Mitarbeit bei der Kontrolle und beim Ausbau der Metalemmaliste
 - durch Integration der eigenen Wörterbuchlemmaliste
 - durch Bestätigung, Ablehnung oder Kommentierung einer vorhandenen (möglicherweise automatisch erzeugten) Lemmazuordnung
 - Vernetzung der eigenen Wörterbuchartikel mit allen angeschlossenen Wörterbüchern über die Metalemmaliste
 - Verfügbarmachung der eigenen Ergebnisse für die Science Community



Herzlichen Dank für Ihre Aufmerksamkeit!



Andrea Rapp, Kompetenzzentrum Trier

www.kompetenzzentrum.uni-trier.de

andrea.rapp@uni-trier.de

