
Beteiligte Projekte

- Ausbau und Pflege der Korpora geschriebener Sprache
- Methoden der Korpusanalyse und -erschließung
- Wechselwirkungen zwischen linguistischen und bioinformatischen Verfahren, Methoden und Algorithmen

Beteiligte Mitarbeiter

- Dipl.-Ing. Cyril Belica
- Dr. Marc Kupietz
- Dr. Harald Längen
- Dipl.-Inf. Rainer Perkuhn
- Heike Stadler
- Dr. Andreas Witt

Kontakt

Dipl.-Ing. Cyril Belica
Programmbereich Korpuslinguistik
Institut für Deutsche Sprache
Postfach 10 16 21
D-68016 Mannheim

Telefon: 0621/1581-0
Fax: 0621/1581-200
E-Mail: korpuslinguistik@ids-mannheim.de



 INSTITUT FÜR
DEUTSCHE SPRACHE

R 5, 6-13

D-68161 Mannheim

www.ids-mannheim.de

Mitglied der

Leibniz-Gemeinschaft

Das Institut für Deutsche Sprache (IDS) ist die zentrale Einrichtung zur Erforschung und Dokumentation der deutschen Sprache in ihrem Gebrauch und in ihrer neueren Geschichte. Es gehört zu den 86 außeruniversitären Forschungseinrichtungen und Serviceeinrichtungen für die Forschung der Leibniz-Gemeinschaft.

 INSTITUT FÜR
DEUTSCHE SPRACHE

Programmbereich Korpuslinguistik

CCDB

***Eine korpuslinguistische
Denk- und
Experimentierplattform***

Um linguistische Phänomene konsequent empirisch erschließen zu können, ist es oft notwendig, direkten Zugang zu enormen Datenmengen zu haben, die von großen Korpora abgeleitet wurden. Die Kookkurrenzdatenbank CCDB ist eine korpuslinguistische Denk- und Experimentierplattform, ein *transparentes Labor* zur Erforschung, Entwicklung und Evaluierung einer Methodik zur datengeleiteten Exploration und Modellierung des Sprachgebrauchs.

Die aktuelle Version der CCDB

- dient projektintern als empirische Grundlage für die Formulierung von neuen sprachgebrauchsbezogenen linguistischen Hypothesen
- integriert Early-Beta-Implementationen von Ergebnissen laufender Forschungsarbeiten: derzeit u.a. die Modellierung semantischer Verwandtschaft oder die Visualisierung von Verwendungsaspekten
- basiert auf einem Ausschnitt von über 2,2 Milliarden Textwörtern aus dem DEUTSCHEN REFERENZKORPUS (DEREKO)
- umfasst reichhaltige Kookkurrenzprofile für ca. 220.000 Bezugswörter, die syntagmatische Muster und andere Informationen zum Gebrauch dieser Wörter beinhalten
- enthält für jedes Bezugswort bis zu fünf verschiedene Kookkurrenzprofile mit jeweils bis zu 100.000 KWIC-Kontexten
- stellt ein Netz von Ähnlichkeitsrelationen zwischen Kookkurrenzprofilen bereit
- ist eingeschränkt öffentlich zugänglich unter <http://corpora.ids-mannheim.de/ccdb/>

Analysewort: **vergewissern**, Analysetyp 0

+ 3 5 355	wirklich auch	21	71%	sich zu vergewissern ob daß dass ... auch wir
+ 3 5 355	wirklich	63	80%	sich zu vergewissern ob daß dass ... auch wir
+ -4 -2 289	sollte vorher auch	1	100%	sollte ... vorher vergewissern ... auch
+ -4 -2 289	sollte vorher	9	88%	sollte [man sich] vorher [...] vergewissern ob ...
+ -4 -2 289	sollte auch	15	86%	sollte man sich ... vergewissern daß dass ob c
+ -4 -2 289	sollte Fahrzeug	3	66%	sollte man sich vergewissern ... das Fahrzeug
+ -4 -2 289	sollte	71	92%	sollte man sich ... vergewissern ob daß der
+ -3 -2 271	eigenen Existenz	6	83%	der eigenen Existenz [zu] vergewissern
+ -3 -2 271	eigenen Geschichte	4	100%	seiner eigenen Geschichte [zu] vergewissern
+ -3 -2 271	eigenen Augen	3	100%	mit eigenen Augen [zu] vergewissern
+ -3 -2 271	eigenen	66	77%	sich der eigenen ... zu vergewissern und

Modelling Semantic Proximity, version 0.21, Hierarchical Clustering

↑ Weichsel ↓
click on arrows to expand/collapse

- Kirsche Birne Pfirsich Erdbeere Aprikose Apfel Weintraube Toma
- Landwehrkanal
- Neiße Wolga Grenzfluß Ostpreuße Galizien Flüsschen Landstrich Norden Litauen Baltikum Partisane Hügel Bergregion
- Elbe Jangtse Nebenfluß Fluss Fluß Oberlauf Hochwasser Wasse

Modelling Semantic Proximity, version 0.19, Self-Organizing Map (initial tau: 0.02, distance: uncentered Pearson, iterations: 10)

Identität
click on a cell to toggle label locking

Kulturerbe Institution Hegemonie nationalstaatlich supranational Abgrenzung Staatlichkeit territorial	Rechte Zusammenhalt Angelegenheit Machtstruktur Struktur Errungenschaft Friedensordnung Einigungsprozeß	Eigenständigkeit Souveränität Integrität Handlungsfähigkeit Autonomie Besitzstand Bedeutung Entscheidungskompeten	Selbstständigkeit Kompetenz Gleichgewicht Selbstständigkeit Verantwortlichkeit Lebensgrundlage Glaubwürdigkeit Existenzgrundlage	Todesursache Tathergang Unschuld Staatsangehö Intimsphäre Vaterschaft Aufenthaltsort Identifizierung
Selbstbestimmung national herausbilden Gesinnung Vereinnahmung Repräsentation Pluralismus Schicksalsgemeinschaft	Erbe staatsbürgerlich Unterordnung	Selbstachtung Freiheit Zusammengehörigkeitsg	Würde unwiederbringlich	Individualität Inimität Schuld Authentizität Schicksal Rückschluß Geborgenheit Lebenssinn
Zugehörigkeit ethnisch rassisch Chauvinismus Minderheit Zionismus Überfremdung Bewahrung	Selbstverständnis Selbstbehauptung Heimstatt dörflich	Zusammengehörigkeit Emanzipation kulturell Bewußtsein Selbstverwirklichung Pluralität Liberalität Urbanität	Orientierung Einzigartigkeit Identifikation Eindeutigkeit Einmaligkeit Ursprünglichkeit Leben Lebendigkeit	Potenz Profil Befriedigung Gestaltungskr
religios Assimilation Entwurzelung Wiedergeburt Auslöschung Selbstvergewisserung völkisch Unterwerfung	Rückbesinnung Identitätsfindung Verkörperung	Verwurzelung Verschiedenheit Gemeinsamkeit Selbstbild Überbau Eigensinn	Bindung Bezugspunkt Homogenität Betätigung Zerrissenheit Subjektivität	Charakter Persönlichkeit Perspektive Dimension Bezug Gedächtnis Umfeld Psyche
Herkunft Tradition Religion Weltanschauung Musikkultur Religiosität Kultur Schmelztiegel	Männlichkeit Menschsein geschlechtlich Weiblichkeit Verstümmelung Fremdheit Sexualität Körperlichkeit	Vielfalt Ausdrucksform Geschlecht spezifisch Lebensform Lebenswelt genuin Vorstellungswelt	Sozialisation Ausprägung Prägung Lebensentwurf Betonung Heterogenität Empfindung widerspiegeln	Eigenheit Eigenart Besonderheit Präferenz Befindlichkeit Zuordnung Rücksicht Ausdruckswei

Stand: 04/12

Textwörter passend zum Suchmuster
Atriffig treten als Kookkurrenzpartner bei
folgenden Analysewörtern und -typen auf

Wort	0	1	2	3	4
Grund					
triffig					
Grund					
triffige					
Grund					
triffiger					
Argument					
triffige					
Entschuldigungsg					
triffiger					
Scheidungsgrund					
triffige					
Asylgrund					
triffigen					
Gegenargument					
triffige					
vorliegen					
triffige					
fernbleiben					
triffigen					
vorbringen					
triffige					
anführen					
triffige					
fernbleiben					
triffiger					
Vorliegen					
triffiger					
vorbringen					
triffigen					
Folgende verwandte Kookkurrenzprofile zu Charakteristik wurden gefunden (anklickbar, absteigend nach Verwandtschaftsgrad sortiert):					
Merkmal					
Eigenheit					
Ausprägung					
Element					
Eigenart					
Parameter					
Eckpfeiler					
Charakterzug					
Eigenschaft					
Ausdrucksform					
Bezugspunkt					
Ausformung					
Anliegen					
Stilmittel					
Charakteristik					
Besonderheit					
Charakter					
Kriterium					
Ausdrucksmittel					
Affinität					