



ANNIS Einführung RIDGES Herbology 5.0-beta1

Laura Perlitz und Carolin Odebrecht Humboldt-Universität zu Berlin



Korpus RIDGES Herbology



- Wiederholung
 - Kräuterkundekorpus
 - 15.-20. Jahrhundert
 - verschiedene Dialekte aus dem deutschsprachigen Raum
 - Arbeit von vier Seminaren aus B.A. und M.A. von ca. 60 Studenten
 - Projekthomepage
 - http://korpling.org/ridges/
 - Annotationsrichtlinie
 - http://korpling.org/ridges/documentation_de/

ANNIS



Such- und Visualisierungstool

- webbasiertes Suchtool f
 ür Korpora
 - Projekthomepage: <u>http://corpus-tools.org/annis/</u>
 - Zugang auf das ANNIS-Tool: <u>https://korpling.org/annis3</u>
- Was kann ANNIS?
 - findet Annotationen in einem Korpus
 - findet mehrere Annotationen in bestimmter Relation zu einander
 - gibt die gesuchten Annotationen als Treffer in vielfältigen Visualisierungen aus
 - filtert nach Metadaten
 - exportiert Treffer
 - u.v.m.
- Hier im Seminar lernen wir mit dem Korpus RIDGES Version 5.0beta1, wie man ANNIS benutzt!

Weitere Informationen zu ANNIS

- Zum Nachlesen:
 - Dokumentation: <u>http://corpus-tools.org/annis/documentation.html</u>
 - Nutzerhandbuch: <u>http://corpus-</u> tools.org/annis/resources/ANNIS_User_Guide_3.4.3.pdf
 - Krause, Thomas and Zeldes, Amir (2016) "ANNIS3: A New Architecture for Generic Corpus Query and Visualization". *Digital Scholarship in the Humanities* 31(1), 118-139. <u>http://dsh.oxfordjournals.org/cgi/content/abstract/fqu057?ijkey=GJBr</u> <u>OLhNfKW1g8i&keytype=ref</u>
- Zum Nachschauen (Video-Tutorials):
 - <u>https://www.linguistik.hu-</u> berlin.de/de/institut/professuren/korpuslinguistik/corpus-tools/annistutorials</u>



Historische Korpora Variation, Segmentierung

Historische Texte



- Varianz der Orthographie bzw. Setzung
- vieles nicht wirklich "vorhersehbar"
- graphische/(text-) strukturelle Informationen
 - Zeilenumbruch, Seitenumbruch, Setzung



Normalisierung

- Um alle Schreibvarianten eines Wortes in historischen Korpora finden zu können, sind Normalisierungen notwendig.
- In RIDGES gibt es daher drei Textebenen:
 - **dipl**: diplomatische Transkription
 - clean: Ersetzung aller heute unüblichen Sonderzeichen und Zusammenziehung von Wörtern, die durch einen Zeilenumbruch getrennt wurden (mit Bindestrich)
 - **norm**: Anpassung an die moderne Orthographie

Segmentierung I

- Annotationen in RIDGES (vgl. Belz et al. 2015)
 - dipl: diplomatische Transkription
 - clean: automatische Normalisierungen
 - norm: manuelle Normalisierung
 - lb: Zeilenumbrüche
 - lemma: Lemmatisierungen

clean	die	aus	drei	Worten	zusammenge	esetzt	sind
dipl	die	aus	drei	Worten	zulammen∞	geletzt	lind
lb	lb					lb	
lemma	d	aus	drei	Wort	zusammense	etzen	sein
norm	die	aus	drei	Worten	zusammenge	esetzt	sind

Trefferreferenz: https://korpling.org/annis3/?id=cb133ed8-ae93-4132-82d3-642ef9b92f58

Belz, Malte; Odebrecht, Carolin; Perlitz, Laura & Voigt, Vivian (2015): Annotationsrichtlinien zu Ridges Herbology Version 4.1, Humboldt-Universität zu Berlin.

Segmentierung II

- Segmentierung
 - mehrere primäre Textebenen in RIDGES: dipl, clean, norm
 - Jede dieser Ebenen besitzt eine eigenständige Einteilung in kleinste Einheiten (Segmentierung)
 - Bei der weiteren Annotation muss IMMER festgelegt werden, auf welcher Segmentierung sie sich bezieht!

	clean	die	aus	drei	Worten	zusammenge	esetzt	sind
1	dipl	die	aus	drei	Worten	zulammen∞	geletzt	lind
	lb	lb					lb	
^	lemma	d	aus	drei	Wort	zusammense	etzen	sein
6	norm	die	aus	drei	Worten	zusammenge	esetzt	sind

Trefferreferenz: https://korpling.org/annis3/?id=cb133ed8-ae93-4132-82d3-642ef9b92f58

Segmentierung III

Position dipl	1	2	3		4		5		6	7	Anzahl der
Position norm	1	2	3		4		5			6	Einheiten
clean	die	aus	drei		Wor	ten	zus	sammenge	setzt	sind	dipl=7 norm=6
dipl	die	aus	drei		Wor	ten	zul	ammen∞	geletzt	ſind	
lb	lb								lb		
lemma	d	aus	drei		Woi	t	zu	sammense	etzen	sein	
norm	die	aus	drei		Woi	rten	zu	sammenge	esetzt	sind	
Position dipl	1	2		3		4		5			 Anzahl der
Position dipl Position norm	1	2 2		3 3		4 4		5 5	6		Anzahl der Einheiten
Position dipl Position norm clean	1 1 Den	2 2 Winter-s	pinet	3 3 Se	hr	4 4 gro	SZ	5 5 zubekon	6 nmen		Anzahl der Einheiten dipl=5 norm=6
Position dipl Position norm clean dipl	1 1 Den Den	2 2 Winter-s Wintersfp	pinet	3 3 Se	ehr hr	4 4 gro gro	sz íz	5 5 zubekon zubekon	6 nmen nmen		Anzahl der Einheiten dipl=5 norm=6
Position dipl Position norm clean dipl lb	1 1 Den Den Ib	2 2 Winter-s Winter₅fp	pinet pinet	3 3 Se	ehr hr	4 4 gro gro	sz íz	5 5 zubekon zubekon	6 nmen nmen		Anzahl der Einheiten dipl=5 norm=6
Position dipl Position norm clean dipl Ib Iemma	1 1 Den Den Ib d	2 2 Winter-s Winter₅fp Wintersp	pinet binet	3 3 fel	ehr hr	4 4 gro gro gro	sz íz	5 5 zubekon zubekon zu	6 hmen hmen bekomr	nen	Anzahl der Einheiten dipl=5 norm=6

Trefferreferenz 1: https://korpling.org/annis3/?id=cb133ed8-ae93-4132-82d3-642ef9b92f58

Trefferreferenz 2: https://korpling.org/annis3/?id=4c675f63-5890-420e-b3a7-fd9e637b31a3



Suche

ANNIS Tutorial mit historischen Korpora (RIDGES Herbology)

Bitte öffnen!

- RIDGES-Dokumentation (Version 4.1, für den Kurs gibt es eine aktualisierte Annotationsrichtlinie): <u>http://korpling.org/ridges/documentation_de/</u>
- STTS: <u>http://www.ims.uni-</u> <u>stuttgart.de/forschung/ressourcen/lexika/TagSets/stts-</u> <u>table.html</u>
- ANNIS-Instanz: https://korpling.org/annis3/

https://korpling.org/annis3/

■< 🕅 About ANNIS		Report Pro	oblem				<u>Help us to ma</u>	ke ANNIS betterl	not logged in 🔒 Log	gin		
Please enter A01	. mie	ru				• Help/Examples						
fittabe enote ng.	u que	1				<u>Tutorial</u>						
				Quer	y or	🔳 Example Querie	25					
				Duild		Example Query	Example Query Description open corpus brow					
			L		#	Q pos="NN"		Findet einen Nomen.	RIDGES Herbology Version4.1	^		
						Q lang="foreign	<u>"</u>	search for foreign language (non-Hausa) segments within the Hausa text	a5.hausa.umarnin.uwa V2			
						Q pos="NPR	Die Version 5 ()-heta1 ist derzeit nocl	n nicht	Ш		
						Q <u>norm="אסא</u>						
Q Search Mor	e 🔻	History	•			Q pos="VSTA	offentlich zuga	anglich. Daher ist eine	<u>.m.11</u>	_		
Welcome to ANNISI A to on the right side.	utorial	is availabl	e			Q <u>norm="৯π</u>	Anmeldung er	forderlich. Für den Kur	s wurde m.11			
on the right side.					Q <u>morph="ա</u>	۹ morph="weight folgendes Konto angelegt:						
Corpus List Search O	ptions					Q pos="PPER	Q pos="PPER					
Visible: All				•][2	Q norm="NKC				-		
Filter						O edition="er	Renutzer hie	stkorn		-		
Name -	lexts	Tokens						bikorp		_		
a5.hausa.umarnin.uwa_\	47	10.194	0			Q document	Passwort: 52	20049	<u>el</u>	_		
abraham.our.father	7	7.671	0	Ē		Q <u>edition="er</u>		or the calcon				
apophthegmata.patrum.	11	1.500	0	₿		Q <u>document</u>		Access all texts in the corpus for reading purposes	DDD-Genesis			
arabic.tree.test	1	11	0	B		Q <u>edition="enti</u>	<u>.</u>	Search for the form "enti" in the transcription of the edition	DDD-Heliand			
b2.hausa	50	6.991	0	Ē		Q <u>document</u>		Access all texts in the corpus for reading purposes	DDD-Heliand			
b4.tatian2.0	2.031	11.295	0	Ē		Q <u>edition="enti</u>		Search for the form "enti" in the transcription of the edition	DDD-Isidor			
b4.tatian2.1	2.030	11.295	0	Ē		Q document		Access all texts in the corpus for reading purposes	DDD-Isidor			
b7.wolof.web.V2	4	14.676	0			Q <u>edition="enti</u>		Search for the form "enti" in the transcription of the edition	DDD-IsidorLatein			
b7.wolof.wiki.V4	14	12.737	0			Q document		Access all texts in the corpus for reading purposes	DDD-IsidorLatein	-		
BeMaTaC 1 2013-01	12	11.192	6					harbarr		_		

ANNIS Tutorial mit historischen Korpora (RIDGES Herbology)



Interface

Startseite, Hilfestellung, Query Fenster

ANNIS Query Language



- oder: Wie sage ich ANNIS, was ich suche?
- für die Suche nach Annotationen gelten zwei Prinzipien:
 - Attribut-Wert-Paar (AW-Paar)
 - Relationen zwischen AW-Paaren

Prinzip I: Attribut-Wert-Paar



- Voraussetzung für das Funktionieren der Suche ist das Vorhandensein einer Ebene namens "dipl" (Dokumentation!).
 Erwartetes Ergebnis ist es, exakt alle Vorkommen dieser
 - Zeichenkette in "dipl" im ausgewählten Korpus zu finden.

dipl	wider	auff	ein	friſches	zerítolíenes	kraut	/	thue	es	wider	in
clean	wider	auff	ein	frisches	zerstossenes	kraut	/	thue	es	wider	in
norm	wieder	auf	ein	frisches	zerstoßenes	Kraut	/	Tue	es	wieder	in

Attribut-Wert-Paar dipl=/kraut/



🛋 🙀 About ANNIS 🖂 Report Probler	m	Help us to make ANNIS better!	logged in as "laura" 🔒 Logout
		Help/Examples Q. Query Result ×	
dipl=/kraut/	2	Base text v	
	Query	[K] 1 / 26 [X] Displaying Results 1 - 10 of 258	Result for: dipl=/kraut/
	Builder	A Path: RIDGES Herbology Version5 O-beta1 > AlchymistischePractic 1603 Libavius (diol 1915 - 1925)	left context: 5 v right context: 5 v
		W. b. 1. geody rets kraut / zernacket es ein wenig	
		the lexit Layers	
		B Switzer (anotation	
		B Granica Amountain	
Q Search More V History V			
258 matches		Other Annotation	
in 15 documents			
		B Diplomatic Text	
Corpus List Search Options		Normalized Text	
Visible: All	v 2	2 👩 < Path: RIDGES_Herbology_Version5.0-beta1 > AlchymistischePractic_1603_Libavius (dipl 2466 - 2476)	left context: 5 🗸 right context: 5 🗸
ridges		diftillirte waller vff ein frilches kraut / vnd diftillir es wider	
Name Texts Tokens		H Text Lavers	
Ridges-SE-CompAnalyse 20 92,368	0 🖹	Lexical Annotation	
ridges herhology 14 62 724	A B	Morphological Annotation	
Thiges.nerbology		Syntactic Annotation	
RIDGES_Herbology_Versior 22 122,698	0	Graphical Annotation	
RIDGES_Herbology_Versior 29 154,266	0	Content Annotation	
RIDGES Herbology Versior 29 154,267	0	Other Annotation	
DIDCEE Lightelegy Versie 26 192 710		All Annotation	
RIDGES_Herbology_version 36 183,710		Diplomatic Text	
Ridges_Herbology_Version 13 60,811	0	Normalized Text	
		3 O < Path: RIDGES_Herbology_Version5.0-beta1 > AlchymistischePractic_1603_Libavius (dipl 3870 - 3880)	left context: 5 🗸 right context: 5 🗸
		wider auff ein frilches zerltollenes kraut / thue es wider in	
		🗄 Text Layers	
		Lexical Annotation	
		Morphological Annotation	
		Syntactic Annotation	
		Graphical Annotation	
		Content Annotation	
		Other Annotation	
		All Annotation	
		B Diplomatic Text	
		Normalized Text	
		4 🚯 < Path: RIDGES_Herbology_Version5.0-beta1 > AlchymistischePractic_1603_Libavius (dipl 4005 - 4015)	left context: 5 🗸 right context: 5 🗸

Suchreferenz: https://korpling.org/annis3/?id=8ebb8273-3048-4c06-b697-81a19573eae4

ANNIS Tutorial mit historischen Korpora (RIDGES Herbology)

Aufgabe Schreibvarianten



- Finden Sie alle Schreib- und Flexionsvarianten von Kraut!
 - Für die Suche nach allen Varianten wählen Sie die passende Annotationsebene (Attribut) aus
 - Passendes Attribut wäre hier lemma.
 - und setzen den Wert /Kraut/ ein.
 - lemma=/Kraut/

Aufgabe

Schreibvarianten



🚍 🛛 🕌 About ANNIS 🛛 🖂 Repo	ort Prob	lem	Help us to make ANNIS better!	logged in as "histkorp" 🔒 Logout
			Help/Examples Q Query Result ×	
lemma=/Kraut/		3	Base text 🗸	
		Query	く (1 / 59)) Displaying Results 1 - 10 of 586	Result for: lemma=/Kraut/ \prec 🗧
		Builder	1 🚯 🔫 Path: RIDGES_Herbology_Version5.0-beta1 > AlchymistischePractic_1603_Libavius (dipl 888 - 898)	left context: 5 🗸 right context: 5 🗸
		******	weilz vnd weg / aulz <mark>kråutern</mark> / blettern / wurz tzeln	
			⊕ Text Layers	
			🗄 Lexical Annotation	
			Morphological Annotation	
			Syntactic Annotation	
Q Search More - His	story 🗸		Graphical Annotation	
	-	_	Content Annotation	
586 matches			Other Annotation	
in 31 documents			All Annotation	
Communities Consult Options			Diplomatic Text	
Corpus List Search Options			Normalized Text	
Visible: All		v (2	2 ❶	left context: 5 🗸 right context: 5 🗸
Filter			weg / viel waffer aufz <mark>Kråutern</mark> vff einmal zu diftilliren .	
Name	Texts	Tokens	🗄 Text Layers	
MHD context	4	2.760	🗄 🗄 Lexical Annotation	
WHD_CONTEXT	4	2,700	Morphological Annotation	
NoSta-D-1.4-bematac	22	25,934	⊞ Syntactic Annotation	
NoSta-D-1.4-falko	10	6,034	Graphical Annotation	
NoSta-D-Anselm	2	2,710	Content Annotation	
NoSta-D-Kafka	2	10,388		
NoSta-D-TueBaDZ	2	10.832	Diplomatic Text	
Note D Heleve	2	14,040	⊞ Normalized Text	
Nosta-D-Onicum	2	11,312	3 👩 👒 Path: RIDGES_Herbology_Version5.0-beta1 > AlchymistischePractic_1603_Libavius (dipl 961 - 971)	left context: 5 🗸 right context: 5 🗸
pcc2	2	399	/ denn nimb was f <mark>ůr Kråuter</mark> du wilt / zerítoíz die	
ridges.herbology	14	63,734	⊕ Text Layers	
RIDGES_Herbology_Version3.0	22	122,698	⊞ Lexical Annotation	
RIDGES_Herbology_Version4.0	29	154,266		
RIDGES_Herbology_Version4.1	29	154,267	Syntactic Annotation	
RIDGES_Herbology_Version5.0-beta1	36	183,710	E Content Annotation	
Ridges_Herbology_Version_2.0	13	60,811	Other Annotation	
SMULTRON Banana	2	3 782	⊕ All Annotation	
4	-	5,702	⊞ Diplomatic Text	

Suchreferenz: https://korpling.org/annis3/?id=61d11111-124f-42e9-8a4d-7306417f9007

Normalisierung



 Schauen Sie sich die Treffer zu lemma=/Kraut/ genau an!

1 🚯	< Path	: RIDGE	S_Herbo	ogy_	Version	5.0-beta1 >	Alch	iymistische	Prac	tic_1603	3_Libaviu	us (dipl 888	- 898
weiſz	vnd we	eg /	auíz krá	iuter	rn / b	olettern /	wu	r≠ tzeln					
🖯 Tex	xt Layers	-							-				
dipl	weiſz	vnd	weg	1	auíz	kråutern	1	blettern	1	wur∞	tzeln		
clean	weisz	vnd	weg	1	ausz	kräutern	1	blettern	1	wurtze	eln		
norm	Weisen	und	Wege	1	aus	Kräutern	1	Blättern	1	Wurze	eln		

• Es werden alle historischen Wortformen von Kraut gefunden!

Trefferreferenz: https://korpling.org/annis3/?id=2d1e7843-0bf8-45fb-b752-29f8f9415c75

Aufgabe Normalisierung



- Suchen Sie das Lemma von *zusammensetzen*!
 - Finden von ganz unterschiedlichen historischen Schreibvarianten!
 - lemma=/zusammensetzen/
 - Treffer 1: <u>https://korpling.org/annis3/?id=3c470377-a87d-464f-82ce-5e2b7310809c</u>
 , die aus drei Worten zuſammen^s geſetzt ſind , wie die oben
 - Treffer 2: <u>https://korpling.org/annis3/?id=e46cc32b-3f3c-472d-8d2e-e77b2e326c55</u> muſz aus zwei einfachen Worten zu^s ſammengeſetzt ſein , von denen der
 - Treffer 3: <u>https://korpling.org/annis3/?id=83497961-4769-4382-91cc-6af6b6839c4d</u> Wenn der lateinifche Artname zuſammengeſetzt iſt , kann man entweder



Mustersuche Operatoren

ANNIS Tutorial mit historischen Korpora (RIDGES Herbology)

Operatoren für die Mustersuche



Für die Manipulation der **Annotationswerte** zu nutzen:

- . Ein beliebiges Zeichen
- ? 0 oder 1 Zeichen (des vorherigen Elementes)
- * 0 bis unendlich viele Zeichen (d. vorh. E.)
- + 1 bis unendlich viele Zeichen (d. vorh. E.)
- wörtlich (folgendes Zeichen, z.B. spitze Klammern, Punkte)
- (a|b) a oder b (auch: [ab])

Für die Manipulation der **gesamten Anfrage** zu nutzen:

nicht (steht direkt vor dem "=" zwischen Attribut und Wert)

Möchte man nach Unicodes suchen, muss man diese mit einem Backslash \ maskieren, z.B. dipl=/\u002F/ (Hier wird nach "/" gesucht).

Aufgabe Operatoren



- Welche Ergebnisse erwarten Sie für folgende Anfragen, sogenannte Mustersuchen?
 - norm=/g.b./
 - gebe, gibt (für RIDGES, theoretisch noch andere möglich)
 - dipl=/r(a | o)t/
 - rot, rat (für RIDGES)
 - dipl=/meint?/
 - mein, meint, meine (für RIDGES)

Aufgabe Wortarten I



- Wie finden Sie alle Appellativa in RIDGES?
 - Passendes Attribut wäre hier **pos.**
 - pos=/NN/

Aufgabe Wortarten I



E< About ANNIS	🛛 Report Pro	blem	Help us to make ANNIS betteri	logged in as "histkorp" 🔒 Logout
(NN /			Help/Examples Q Query Result ×	
pos=/NN/			Base text v	
		Query	K K 1 / 3363 X Displaying Results 1 - 10 of 33628	Result for: pos=/NN/ < 🌻
		Builder	1 👩 🔫 Path: RIDGES_Herbology_Version5.0-beta1 > AlchymistischePractic_1603_Libavius (dipl 1 - 8)	left context: 5 🗸 right context: 5 🗸
			An den Lefer . ES findt mir lieber	
			⊞ Text Layers	
			⊞ Lexical Annotation	
			Morphological Annotation	
			Syntactic Annotation	
Q Search More	🖌 🗌 History 🗸	•	Graphical Annotation	
22620 metches			H Content Annotation	
in 36 documents			H Other Annotation	
Corpus List Search Opti	ions			
Visible: All		v 2	2 👩 🛃 Path: RIDGES_Herbology_Version5.0-beta1 > AlchymistischePractic_1603_Libavius (dipl 4 - 14)	left context: 5 🗸 right context: 5 🗸
Filter			. ES findt mir lieber Lefer / zwey feine Alchy- miftifche	
Name	 Text 	s Tokens	E Text Layers	
MUD contaut	4	2,760	Lexical Annotation	
MHD_context	4	2,700	Morphological Annotation	
NoSta-D-1.4-bematac	22	25,934	Syntactic Annotation	
NoSta-D-1.4-falko	10	6,034	Graphical Annotation	
NoSta-D-Anselm	2	2,710		
NoSta-D-Kafka	2	10,388		
NoSta-D-TueBaDZ	2	10.832		
NoSta-D-Unicum	2	11 312	Normalized Text	
10500 5 0110011	2	200	3 👩 < Path: RIDGES_Herbology_Version5.0-beta1 > AlchymistischePractic_1603_Libavius (dipl 10 - 20)	left context: 5 🗸 right context: 5 🗸
pcc2	2	399	/ zwey feine Alchy, miltilche Tractåtlein von guten freundten zukomen /	
ridges.herbology	14	63,734	⊞ Text Layers	
RIDGES_Herbology_Version	13.0 22	122,698	Lexical Annotation	
RIDGES_Herbology_Version	14.0 29	154,266	Morphological Annotation	
RIDGES_Herbology_Version	14.1 29	154,267	Syntactic Annotation Graphical Apportation	
RIDGES_Herbology_Version	15.0-beta1 <u>36</u>	183,710		
		60.014	T Other Apportation	

Suchreferenz: https://korpling.org/annis3/?id=11b66562-075d-420b-bae1-8062d076d9c4

Aufgabe Wortarten II



- Wie finden Sie alle Adjektive unabhängig von ihren Bezugswörtern in RIDGES?
 - Was sagen die Richtlinien (STTS)?
 - Passendes Attribut wäre hier **pos.**
 - STTS: Unterschied ADJA und ADJD?
 - pos=/ADJ./

Aufgabe Wortarten II



🛋 🙀 About ANNIS 🛛 🖂 Report Probl	em		Help us to make ANNIS better!		logged in a	as "laura" 🛛	👗 Log	out
		_	Help/Examples Q Query Result ×					
pos=/ADJ./	. 9	2	Base text v					
	Qu	lery	K (1 / 1690) Displaying Results 1 - 10 of 16895		Result fo	r: pos=/ADI	/ <	Т:
	Bui	ilder	A Path: RIDGES_Herbology_Version5,0-beta1 > AlchymistischePractic 1603_Libavius (dipl 3 - 13)	left context:	5 🗸 rig	tht context:	5	4
		****	Lefer . ES findt mir lieber Lefer / zwey feine Alchy			,		-
			Lexical Annotation					
			Morphological Annotation					
			⊞ syntactic Annotation					
O Search Mere Hictory			Graphical Annotation					
			Content Annotation					
16895 matches			Other Annotation					
in 36 documents			All Annotation					
			🗄 Diplomatic Text					
Corpus List Search Options			Normalized Text					
Visible: All	×	2	2 👩 < Path: RIDGES_Herbology_Version5.0-beta1 > AlchymistischePractic_1603_Libavius (dipl 7 - 17)	left context:	5 🗸 rig	ght context:	5	~
ridges			mir lieber Lefer / zwey <mark>feine</mark> Alchy, miltifche Tractâtlein von guten					
Name Texts Tokens			Text Layers					
Ridges-SE-CompAnalyse 20 92,368	0	B	Lexical Annotation					
ridges berbology 14 63.734	A	B	Morphological Annotation					
14 03,734	•	8	Syntactic Annotation					
RIDGES_Herbology_Versior 22 122,698	0		🗄 Graphical Annotation					
RIDGES_Herbology_Versior 29 154,266	0		Content Annotation					
RIDGES_Herbology_Versior 29 154,267	0	e	Other Annotation					
RIDGES Herbology Version 36 183 710	6	B	All Annotation					
	•	•	Diplomatic Text					
Ridges_Herbology_version, 13 60,811	0		H Normalized Text	left seatests				_
			3 6 < Path: RIDGES_Herbology_versions.u-beta1 > Alchymistiscnevractic_1603_Libavius (dipi 8 - 19)	left context:	5 🗸 rig	int context:	5	~
			lieber Lefer / zwey feine Alchy- miltifche Tractåtlein von guten freundten zukomen					
			🕀 Text Layers					
			Lexical Annotation					
			Morphological Annotation					
			Syntactic Annotation					
			Graphical Annotation					
			Content Annotation					
			Other Annotation					
			All Annotation					

Suchreferenz: https://korpling.org/annis3/?id=9c15ac7c-5053-4586-be0e-5d2587b2e18e

Aufgabe



Suchen von Annotationsebenen

- Wie finden Sie heraus, ob es in RIDGES Herbology Annotationen für Überschriften gibt?
 - Wie sieht eine Anfrage aus, um Überschriften zu finden?
 - Passendes Attribut wäre hier head.
 - head=/head/

Aufgabe



Suchen von Annotationsebenen

🖃 🔛 About ANNIS	ort Probl	em	Help us to make ANNIS better!	logged in as "histkorp"	🔺 Logout	Ī
			Help/Examples Q Query Result ×			
head=/head/		2	Base text v			
		Query	K K 1 /75 N Displaying Results 1 - 10 of 743	Result for: head=/head	d/ \prec	ï
		Builder	1 😝 🛃 Path: RIDGES_Herbology_Version5.0-beta1 > AlchymistischePractic_1603_Libavius (dipl 1 - 9) left c	context: 5 🗸 right context:	5 🗸	ļ
			An den Lefer , ES findt mir lieber Lefer			
			E Text Layers			
			Lexical Annotation			
			Morphological Annotation			
			Syntactic Annotation			
Q Search More - His	story 🗸		Graphical Annotation			
			Content Annotation			
743 matches			Other Annotation			
In 54 documents			□ Al Anotation			
Corpus List Search Options			B logradie lext			
Visible: All			e hvormarzen ekk 2. a. – a. Dath: PIDGES Harholomy Varsions 0. hatal > AlrhymistischaPrartic 1603 Libavius (din) 437 - 453) left (context: 5 v right context:	5 20	
Filter						
Name	 Texts 	Tokens	Hannen zum beiten brauchen . IRACIATVS CHYMISTARYM DE DESTILLATIONIBVS & extractionibus . IM unumini / ja in			
		· · · · · ·				
MHD_context	4	2,760				
NoSta-D-1.4-bematac	22	25,934	Syntactic Annotation			
NoSta-D-1.4-falko	10	6,034	Graphical Annotation			
NoSta-D-Anselm	2	2 710	Content Annotation			
	-	2,710	Other Annotation			
NoSta-D-Kafka	2	10,388	E All Annotation			
NoSta-D-TueBaDZ	2	10,832	E Diplomatic Text			
NoSta-D-Unicum	2	11,312	U Normalized Text	contauti 🕞 🔐 right contauti		
pcc2	2	399	5 6 < Path. RIDGES_HEIDDIO89_VEISIONSU-DECAT > AICHYMISCISCHEPHACUC_TODS_LIDAVIDS (UIP) 510 - 555)	tontext. 5 V fight context.		
ridges.herbology	14	63,734	von ditti= iller Glasern . Ditz nennet man einen Kolben / es ley von einem Glatz oder andern Materien / vnd dere	ibe mulz man klein 7 gr	oiz / ne	
RIDGES Herbology Version3.0	22	122.698	E Text Layers			
PIDGES Harbolomy Version4.0	20	154.266	Lexical Annotation			
	25	134,200	Morphological Annotation			
RIDGES_Herbology_Version4.1	29	154,267	Syntactic Annotation			
RIDGES_Herbology_Version5.0-beta1	36	183,710	E Graphical Annotation			
Ridges_Herbology_Version_2.0	13	60,811	B Cohen Annotation			
SMULTRON_Banana	2	3,782				
-		+				
			C Spronace rest			Ŧ

Suchreferenz: https://korpling.org/annis3/?id=03ad8de5-6570-443d-b9b8-39bc391ed7ed

Aufgabe Kombinierte Suche



- Finden Sie die diplomatische Wortform *Wermut,* die in einer Überschrift vorkommt!
 - Passende Attribute wären hier head und dipl.

Prinzip II: Relationen

1 \

		1)	es gibt mentere Arten von
haad-/haad/			Annotationen!
nead=/nead/	AW-Paari	2)	Wissen, wie diese
	Deletion		Annotationen in Beziehung
0	Relation		zu einander stehen können!
dial_/Marmut/	$\Lambda M Daar 2$	3)	Spannen (head) überlappen
upi-/ weimut/	AVV-Padi Z		Einheiten in einer
			Segmentierung (dipl)!

/ ABSYI	NTHIUM	MON- TANUM	. Pontiſcher	Wermut /	ABSYNTHIU	IM PONTICI	JM . Me	eer Weri	mut / ABSYNT	HIUM MARI-		
⊞ Text Layers												
⊕ Lexical Annotation												
⊕ Morphological Annotation												
Syntactic Annotation												
Graphical Annotation												
dipl	1	ABSYNTHIUM	MON-	TANUM	-	Pontifcher	Wermut	1	ABSYNTHIUM	PONTICUM		Meer
lb	lb			lb		lb						
pb	pb											pb
pb_n	305											306
head	head					head						head

Trefferreferenz: <u>https://korpling.org/annis3/?id=d01793c8-457e-4f02-b1dd-6b1c96192bc1</u>

Aufgabe Überlappung I



🛋 🗽 About ANNIS 🛛 🖂 Report Problem	Help us to make ANNIS better!									
head (head)	Help/Examples Q Query Result ×									
o	Base text 🗸									
dipl=/Wermut/ Query	K 1 /1 > > Displaying Results 1 - 7 of 7 Result for: he									
Builder	1 👩 < Path: RIDGES_Herbology_Version5.0-beta1 > Phythologia_1662_Becher (dipl 1 - 9) left cont									
	Wermut / ABSYNTHIUM . Bergwermut / ABSYNTHIUM MON- TANUM									
	⊞ Text Layers									
	⊞ Lexical Annotation									
	Morphological Annotation									
	Syntactic Annotation									
Q Search More V History V	Graphical Annotation									
	Content Annotation									
7 matches	⊕ Other Annotation									
in 2 documents	⊞ All Annotation									
	Diplomatic Text									
Corpus List Search Options	Normalized Text									
Visible: All	2 👩 🤞 Path: RIDGES_Herbology_Version5.0-beta1 > Phythologia_1662_Becher (dipl 6 - 21) left cont									
Filter	/ ABSYNTHIUM MON- TANUM . Pontifcher Wermut / ABSYNTHIUM PONTICUM . Meer Wermut / ABSYNTHIUM MARI-									
Name Texts Tokens	⊞ Text Layers									
	⊕ Lexical Annotation									
MHD_context 4 2,760	Morphological Annotation									
NoSta-D-1.4-bematac 22 25,934	⊞ Syntactic Annotation									
NoSta-D-1.4-falko 10 6,034	Graphical Annotation									

Suchreferenz: https://korpling.org/annis3/?id=6d1e0e1b-f7ad-45cc-b833-27eab37e03f6

Syntax-Highlighting



- pro AW-Paar eine Farbe hier unser Beispiel:
 - Farbe Rot für alle Werte des Attributs head
 - Farbe Lila für alle Werte des Attributs dipl

```
head=/head/
_0_
dipl=/Wermut/
```

- Treffer erhalten genau diese Farben:
 - Mehrere dipl-Einheiten rot, da sie zusammen in einer Überschrift stehen.
 - Eine dipl-Einheit lila, da dieses der gesuchte dipl-Wert . Pontifcher Wermut / ABSYNTHIUM PONTICUM .

Aufgabe Überlappung II



- Suchen Sie ein Verb, das in einer Überschrift vorkommt!
 - Passende Attribute wären **pos** und **head.**
 - Operator <u>o</u>
 - pos=/V.*/ _o_ head=/head/

Aufgabe Überlappung II



🛋 🙀 About ANNIS 🖂 Report Problem	Help us to make ANNIS better!	logged in as "laura" 🔒 Logout
	Help/Examples Q Query Result ×	
pos=/V.*/	Base text v	
Query	K (1 / 16) Displaying Results 1 - 10 of 152	Result for: pos=/V.*/ o head=/head/ <
Builder	A Path: RIDGES Herbology Version5.0-betal > AlchymistischePractic 1603 Libavius (dipl 510 - 555)	left context: 5 v right context: 5 v
	von diffi- lier Offsern. Dift nannet man einen Kolhen / es fey von einem Glaft oder andern Materian / vod	derfelbő mufz man klein (grofz (bi
		Acheloe maiz man kielin / groiz / m
	E Text Layers	
	Exical Annotation	
	Morphological Annotation	
O Search More - History -	Syntactic Annotation	
Q search More V History V	🗄 Graphical Annotation	
152 matches	Content Annotation	
in 14 documents	Other Annotation	
	All Annotation	
Corpus List Search Options	Diplomatic Text	
Visible: All 🗸 🥑	Normalized Text	
ridges	2	left context: 5 🗸 right context: 5 🗸
Name Texts Tokens	von diftiø lier Glåsern . Difz nennet man einen Kolben / es fey von einem Glafz oder andern Materien / vnd	derfelbê mufz man klein / grofz / h
Ridges-SE-CompAnalyse 20 92,368 🚯 📄	< IT Text Lavers	• • •
ridges.herbology 14 63,734 🚯 🖹		
RIDGES Herbology Version 22 122.698	Morphological Annotation	
	Syntactic Annotation	
RIDGES_Herbology_Version 29 154,266	🗄 Graphical Annotation	
RIDGES_Herbology_Versior 29 154,267 🚯 🖹	Content Annotation	
RIDGES_Herbology_Versior 36 183,710 3	Other Annotation	
Ridges Herbology Version 13 60.811	All Annotation	
	Diplomatic Text	
	Normalized Text	
	3 ● <path: ridges_herbology_version5.0-beta1=""> AlchymistischePractic_1603_Libavius (dipl 510 - 555)</path:>	left context: 5 🗸 right context: 5 🗸
	von difti» lier Glåsern . Difz nennet man einen Kolben / es fey von einem Glafz oder andern Materien / vnd	derſelbẽ muſz man klein / groſz / h
		÷.
	H Text Layers	
	E Lexical Annotation	
	Morphological Antiocation	
	Graphical Annotation	

Suchreferenz: https://korpling.org/annis3/?id=6d1e0e1b-f7ad-45cc-b833-27eab37e03f6

Aufgabe Identität I



- Suchen Sie alle normalisierten Formen von dem Vollverb "setzen", die infinit sind.
- Passende Attribute wären norm und pos.
 - Operator _=_
 - norm=/setzen/_=_ pos=/VVINF/

Aufgabe Identität I



🖃 🔿 😹 About ANNIS	m	Help us to make ANNIS better!	logged in as "laura" 🔒 Logout
		Help/Examples Query Result ×	
norm=/setzen/		Base text v	
= pos=/\//TNE/	Query	K C 1 /2 N Displaying Results 1 - 10 of 11	Result for: norm=/setzen/ = nos=////INE/
p03-/ VVIII /	Builder	A Path: RIDGES Herbology Version5 0-beta1 > AlchymistischePractic 1603 Libavius (dipl /18 - /28)	left context: 5 v right context: 5 v
		nicht verdreiten lallen etwas darzu zuletzen 7 vrid wie viel ich	
		Elevical Appetation	
		El Worphological Annotation	
Q Search More V History V			
11 matches			
in 8 documents			
Corpus List Search Options		B Normalized Text	
Visible: All	~ 2	2	left context: 5 🗸 right context: 5 🗸
ridges		feu∉ lung eines Balnei Roris (etzen / welches derwegen ein Balneum	
Name Texts Tokens			
Pidges SE CompApplyso 20 92.269	A D	T Lexical Appotation	
Ridges-st-companiaryse 20 52,508		T Morphological Annotation	
ridges.herbology 14 63,734	0	E Syntactic Annotation	
RIDGES_Herbology_Versior 22 122,698	0	E Graphical Annotation	
RIDGES_Herbology_Versior 29 154,266	0	E Content Annotation	
RIDGES Herbology Version 29 154 267	6	Other Annotation	
		🗄 All Annotation	
RIDGES_Herbology_Versior 36 183,710	i 🖹	🗄 Diplomatic Text	
Ridges_Herbology_Version_13 60,811	0	🕀 Normalized Text	
		3 🚯 🔩 Path: RIDGES_Herbology_Version5.0-beta1 > ArtzneyBuchleinDerKreutter_1532_Tallat (dipl 2248 - 2258)	left context: 5 🗸 right context: 5 🗸
		vnd widder zu dem fewer fetzen vnd wol fchaumen / vnd	
		Text Layers	
		Exical Annotation	
		Morphological Annotation	
		Syntactic Annotation	
		⊞ Graphical Annotation	
		Content Annotation	
		Other Annotation	
		All Annotation	

Suchreferenz: https://korpling.org/annis3/?id=2f10fab7-70a8-4047-861e-48766a301edb

Aufgabe Identität I



Beachte: Am ersten Treffer wird deutlich, wie wichtig mehrere Segmentierungen und Normalisierungen sind. Deshalb muss man sich für eine Suchanfrage genau überlegen, auf welcher Annotationsebene (Attribut) man was genau suchen möchte.

1 🚯	Path: RIDGES_Herbology_Version5.0-beta1 > AlchymistischePractic_1603_Libavius (dipl 418 - 428)														
nicht	nicht verdriellen lallen etwas darzu <mark>zuletzen</mark> / vnd wie viel ich														
🖯 Tex	Text Layers														
dipl	nicht	nt verdriellen		laffen	etwas	darzu	zuſetze	en /	vnd	wie	viel	ich			
clean	nicht	verdries	ssen	lassen	etwas	darzu	zusetz	en (vnd	wie	viel	ich			
norm	nicht	verdrie	3en	lassen	lassen etwas da		zu s	setzen /	und	wie	viel	ich			
🗆 Lex	kical An	notation			•										
norm	r	nicht	vero	drießen	lassen	etwas	dazu	zu	setzen			und	wie	viel	ich
lemma	a r	nicht ve		drießen	lassen	en etwas da		zu	setzen	1		und	wie	viel	ich
pos	PTKNEG		PTKNEG VVINF		VVFIN	PIS	PAV	PTKZU	VVINF	5(KON	KOUS	PIS	PPER
pos_k	pos_klein PTKNG			F	VFIN	Ν	ADV	PTKZU	VINF	ZE	ICHEN	KOORD	PWAV	Ν	N

Trefferreferenz: https://korpling.org/annis3/?id=a0a4ab8c-695d-4523-843c-4037a13c9775

Aufgabe Identität II



- Suchen Sie Partikelverben, die die Partikel auf beinhalten!
 - Passende Attribute wären lemma und pos.
 - Operator _=_
 - pos=/VV.*/ _=_ lemma=/auf.*/

Aufgabe Identität II



🚍 🛛 🙀 About ANNIS	ort Prob	lem	Help us to make ANNIS better!	logged in as "histkorp" 🛛 🚨 Logout
			Help/Examples Q Query Result ×	
pos=/vv.*/			Base text 🗸	
lemma=/auf.*/		Query	IK / 13 > > >>	Result for: pos=/VV.*/ _=_ lemma=/auf.*/ 🔩 🌻
		Builder	1 👩 < Path: RIDGES_Herbology_Version5.0-beta1 > AlchymistischePractic_1603_Libavius (dipl 1350 - 1361)	left context: 5 🗸 right context: 5 🗸
			herůber lind / mulz man <mark>auff hôren</mark> / das Fewer lallen abgehen	
			⊞ Text Layers	
			E Lexical Annotation	
			I Morphological Annotation	
			Syntactic Annotation	
Q Search More V Hist	torv 🗸		⊞ Graphical Annotation	
	, .		⊕ Content Annotation	
127 matches			Other Annotation	
in 32 documents			E All Annotation	
			🗄 Diplomatic Text	
Corpus List Search Options			Normalized Text	
Visible: All		℃	2 👩 🖪 Path: RIDGES_Herbology_Version5.0-beta1 > AlchymistischePractic_1603_Libavius (dipl 1469 - 1479)	left context: 5 🗸 right context: 5 🗸
Filter	_		heråber fein / zu diftilliren <mark>auffgehôrt</mark> werden . Es were dann	
Name	Texts	Tokens	⊞ Text Layers	
MHD context	4	2 760	Lexical Annotation	
WIND_CONTEXT	4	2,700	Morphological Annotation	
NoSta-D-1.4-bematac	22	25,934	Syntactic Annotation	
NoSta-D-1.4-falko	10	6,034	Graphical Annotation	
NoSta-D-Anselm	2	2,710	Content Annotation	
NoSta-D-Kafka	2	10 388	Other Annotation	
Nosta-D-Kaika	-	10,500	H All Annotation	
NoSta-D-TueBaDZ	2	10,832	B Diplomatic lext	
NoSta-D-Unicum	2	11,312	Normalized Text	loft contaxt: E and right contaxt: E and
pcc2	2	399		
ridges.herbology	14	63,734	wûrd / lit es zeit auffzuhôren / vnd das Fewer lallen	
RIDGES Herbology Version3.0	22	122,698	Exclayers	
RIDGES Herbology Version4.0	29	154 266	Morphological Annotation	
	20	101,200	Syntactic Annotation	
RIDGES_Herbology_version4.1	29	154,267	Graphical Annotation	
RIDGES_Herbology_Version5.0-beta1	36	183,710	⊕ Content Annotation	
Ridges_Herbology_Version_2.0	13	60,811	Other Annotation	
SMULTRON_Banana	2	3,782	All Annotation	
-		· · · · ·	Diplomatic Text	

Aufgabe Identität II



- Suchen Sie Partikelverben, die die Partikel auf beinhalten!
 - passend wären lemma und pos
 - Operator _=_
 - pos=/VV.*/ _=_ lemma=/auf.*/
 - Was findet man damit nicht?

Aufgabe Direkte Präzedenz



- Suchen Sie einen Artikel, der ein Nomen direkt präzediert!
 - Passendes Attribut wäre pos.
 - Operator .
 - pos=/ART/ . pos=/N./
 - Achtung! Hier wird der Punkt-Operator in zwei verschiedenen Funktionen verwendet – sowohl für die Manipulation der Annotationswerte als auch als Relationsoperator zwischen Attribut-Wert-Paaren.

Aufgabe Direkte Präzedenz



🚍 🛛 🙀 About ANNIS 🛛 🖂 Repo	ort Probl	em	Help us to make ANNIS better!	logged in as "histkorp" 🚺 🛔 Logou								
(1971)			Help/Examples Q Query Result ×									
pos=/ART/		2	Base text 🗸									
pos=/N./		Query	Image: Constraint of the second sec	Result for: pos=/ART/ . pos=/N./ 🔩								
		Builder	1 🚯 < Path: RIDGES_Herbology_Version5.0-beta1 > AlchymistischePractic_1603_Libavius (dipl 1 - 8)	left context: 5 🗸 right context: 5 🗸								
			An den Lefer . ES findt mir lieber									
			Text Layers									
			Lexical Annotation									
			Morphological Annotation									
			Syntactic Annotation									
Q Search More - Hist	tory 🗸		🗄 Graphical Annotation									
			Content Annotation									
11113 matches			Other Annotation									
in 36 documents			All Annotation									
Consultint Consult Options			Diplomatic Text									
Corpus List Search Options			Normalized Text									
Visible: All		<u> </u>	2	left context: 5 🗸 right context: 5 🗸								
Filter			findeſt . Ich hab mir das judicium Athenienſium von Hippocrate belieben laſſen									
Name	Texts	Tokens	🕀 Text Layers									
MHD context	4	2 760	Lexical Annotation									
	-	2,700	Morphological Annotation									
NoSta-D-1.4-bematac	22	25,934	B Syntactic Annotation									
NoSta-D-1.4-falko	10	6,034	U Graphical Annotation									
NoSta-D-Anselm	2	2,710	B other fanotation									
NoSta-D-Kafka	2	10,388	B Uther Annotation									
NoSta D TuoPaD7	2	10.922										
NOSta-D-TUEBADZ	2	10,852										
NoSta-D-Unicum	2	11,312	3 A Z Path: RIDGES Herbology Version5.0-beta1 > AlchymistischePractic 1603 Libavius (dipl 115 - 126)	left context: 5 🗸 right context: 5 🗸								
pcc2	2	399	The Getter abr. / und dof. Nochtan (anderlich Dauticher Nation – auto-									
ridges.herbology	14	63,734	Text Lavers									
RIDGES_Herbology_Version3.0	22	122,698	Lexical Annotation									
RIDGES_Herbology_Version4.0	29	154,266	Morphological Annotation									
RIDGES Herbology Version4.1	29	154,267	Syntactic Annotation									
	26	192 710	U Graphical Annotation									
RIDGES_Herbology_version5.0-beta1	36	183,710	H Other Amotation									
Ridges_Herbology_Version_2.0	13	60,811	D Other Annotation									

Suchreferenz: https://korpling.org/annis3/?id=d91c7647-4bc2-4351-b91f-ca2f01df90a6

ANNIS Tutorial mit historischen Korpora (RIDGES Herbology)

Aufgabe Metadaten



- Finden Sie heraus, ob die normierte Wortform Blume in einem Text aus dem Jahr 1543 zu finden ist!
- Wenn ja, wie oft?
 - Passende Attribute wären norm und meta::datum.
 - Verknüpfungszeichen &
 - norm=/Blume/ & meta::datum=/1543/
 - 2 Treffer

Aufgabe Metadaten



🚍 🛛 🙀 About ANNIS 🛛 🖂 Repo	ort Prob	lem								Hel	<u>p us to m</u>	ake ANNIS	better!					1	ogged in as	s "histkorp"	💧 Loj	gout
			0 н	elp/Exam	ples	Q Que	ry Result	×														
norm=/Blume/			Bas	Base text ~																		
meta::datum=/1543/	meta::datum=/1543/ Query						K ✓ 1 ✓										ult for: no	rm=/Blum	e/ & meta:::	datum=/15	43/ 🗳	÷
		Builder	1 0	< Pa	ath: RI	DGES_H	erbology	_Version5	5.0-be	eta1 > New	Kreuterb	uch_1543_	Fuchs (dipl	1897 - 1907)			left contex	kt: 5 🗸	right conf	text: 5	~
			J. jr	n ei≉ n	her y	eden t	olůmen	feind d'	ie d	írey eríte	bletter											
			ΘT	ext Layers	5																	
			dipl	. Jr	n ei≠	ner	yeden	blůmer	n fe	eind die	drey	erfte	bletter									
			clear	i . Jr	i eir	ier	yeden	blumer	n s	eind die	drey	erste	bletter									
			norm	i . In	ein	ier	jeden	Blume	s	aind die	drei	ersten	Blätter									
Q Search More - Hist	tory 🗸		± L	axical Ann	notation	n																
2 matches			⊞ N	orpholog	gical An	notation	1															
in 1 document			E 0	/ntactic A raphical /	.nnotau Annota	tion																
Corpus List Search Options			⊞ c	ontent Ar	notati	on																
Visible: All		~ 2] ⊞ 0	ther Anno	otation																	
Filter			∫ ⊞ A	l Annotat	tion																	
Name	Texts	Tokens		ormalized	d Text																	
MHD_context	4	2,760	2 0	🔫 Pa	ath: RII	DGES_H	erbology	_Version5	5.0-be	ata1 > New	Kreuterb	uch_1543_0	Fuchs (dipl	3837 - 3847)			left contex	t: [5 ♥	right cont	text: 5	•
NoSta-D-1.4-bematac	22	25,934	fpor	<u>1 / da</u>	arumb	das	feine b	lům eine	em	ritterſporn	nit vi	ngleich ift										
NoSta-D-1.4-falko	10	6,034	dinl	fnorn		/ d:	arumh	das fr	eine	einem ritterfoorn nit yngleich ift												
NoSta-D-Anselm	2	2,710	clear	Ritters	sporn	/ di	arumb	das s	seine	blum	einem	ritterspor	n nit	vngleich	ist							
NoSta-D-Kafka	2	10,388	norm	Ritters	sporn	/ da	arum	dass s	seine	Blume	einem	Ritterspo	rn nicht	ungleich	ist							
NoSta-D-TueBaDZ	2	10,832	⊞ L	exical Ann	notation	n																
NoSta-D-Unicum	2	11,312	⊞ N	lorpholog	gical An	notation	i -															
pcc2	2	399	⊞ Sj	/ntactic A	.nnotati	ion																
ridges.herbology	14	63,734	⊞ G	ontent Ar	nnotati	an an																
RIDGES_Herbology_Version3.0	22	122,698	⊞ c	ther Anno	otation																	
RIDGES_Herbology_Version4.0	29	154,266	E A	ll Annotat	tion																	
RIDGES_Herbology_Version4.1	29	154,267	⊞ D	ormalizer	d Text																	
RIDGES_Herbology_Version5.0-beta1	36	183,710																				
Ridges_Herbology_Version_2.0	13	60,811																				
SMULTRON_Banana	2	3,782 🖕																				
•		•																				

Suchreferenz: https://korpling.org/annis3/?id=932395db-fa1e-44a8-b993-4aa86c292221

Tipps



- Die Trefferliste zeigt im Kontext nur die Annotationsebenen (Attribute) an, die auch genau in diesem Trefferkontext annotiert worden sind!
 - Sie können den linken und rechten Kontext immer manuell einstellen, also vergrößern oder verkleinern!
- Schauen Sie in die Korpusmetadaten und Annotationsrichtlinien, um zu wissen, welche Annotationsebenen im Korpus vorhanden sind!
- Nicht alle Dokumente in einem Korpus müssen die gleichen (Anzahl und Typ) Annotationsebenen besitzen!

Zusammenfassung



- Suche in ANNIS basiert auf
 - Attributen (Annotationsebenen) und Werten (Kategorien in den Annotationsebenen), z.B.:
 - Suche von **exakten Werten**, z.B. pos=/ADJA/
 - Suche von **Mustern**, z.B. pos=/ADJ./
 - Attribut-Wert-Paare könne in Relation zu einander gesucht werden, z.B.:
 - Annotationen **überlappen** sich
 - Annotationen decken einen identischen Bereich ab
 - Attribut-Wert-Paare können in Verbindung mit Metadaten gesucht werden, z.B.:
 - pos=/ADJA/ & meta::datum=/1870/



Vielen Dank!

ANNIS Tutorial mit historischen Korpora (RIDGES Herbology)

wichtige **Operatoren** für Mustersuche (AW-Paar)



- Ein beliebiges Zeichen
- ? 0 oder 1 Zeichen (des vorherigen Elementes)
- * 0 bis unendlich viele Zeichen (d. vorh. E.)
- + 1 bis unendlich viele Zeichen (d. vorh. E.)
- \\ wörtlich (folgendes Zeichen)
- ! nicht (steht direkt vor dem "=" zwischen Attribut und Wert)
- [abc] Menge (oder [^abc]=alles außer abc)
- (a|b) a oder b (auch: [ab])
- a<mark>{2,3}</mark> a 2 bis 3mal

wichtige **Relationen** zwischen AW-Paaren



- Direkte Präzedenz
- * Indirekte Präzedenz mit beliebigem Abstand
- .1,2 Entweder direkte Präzedenz oder ein Token dazwischen (beliebig erhöhbar, z.B. .1,4 oder auch .2,8)
- _=_ Identische Abdeckung
- _i_ Inklusion
- _o_ Überlappung
- _I_ linksseitige Überlappung
- _r_ rechtsseitige Überlappung

Übungsaufgaben in ANNIS

• Nach ein paar Aufgaben für Sie zum Üben!



 Suchen Sie technische Begriffe und Erklärungen, die ein getrennt geschriebenes Kompositum beinhalten!



- Suchen Sie technische Begriffe und Erklärungen, die ein getrennt geschriebenes Kompositum beinhalten!
 - Passende Attribute wären term, komp und komp_orth.



- Suchen Sie technische Begriffe und Erklärungen, die ein getrennt geschriebenes Kompositum beinhalten!
 - Passende Attribute wären term, komp und komp_orth.
 - Operatoren **_o_** _=_

- term=/t/ _o_ komp=/k/ _=_ komp_orth=/gtr/



E< About ANNI	S E	Report Pr	oblem	m Help us to make ANNIS better! logged in as "histkor												
term=/t/				1	Help/Examples Q Query Result ×											
° komp=/k/=_ kom	h=/gtr/		Que	l ry ier	Base text K I J J Displaying Results 1 - 10 of 25 Result for: term=/t/_o_komp=/k/_=_komp_orth=/gtr/											
					翻	1 ● Path: RIDGES_Herbology_Version5.0-beta1 > left context: 5 ∨ right context: 5 ∨ AlchymistischePractic_1603_Libavius (dipl 681 - 692) ding zu Afchen brennet / Reuerberir Ofen / darin alle ding mit Path: Ride and the second se										
Q Search Mo 25 matches in 6 documents	ore 🗸	History	•			Exclayers Lexical Annotation Morphological Annotation Syntactic Annotation Graphical Annotation Content Annotation Other Annotation All Annotation Delta Annotation										
Visible: All	options				0	Normalized Text										
Filter				<u> </u>	~	2 🗿 < Path: RIDGES_Herbology_Version5.0-beta1 > left context: 5 🗸 right context: 5 🗸										
Name	• Texts	Tokens			T	AlchymistischePractic_1603_Libavius (dipl 2981 - 2992)										
Maerchenkorpus	211	295.880	0	Ð	^	denn lalle den Buttner ein Eychen Fafzlein machen / mit einem Boden										
Mercurius	2	187.423	0			Lexical Annotation										
MHD_context	4	2.760	0	8		Morphological Annotation										
NoSta-D-1.4-bematac	22	25.934	0			Syntactic Annotation Graphical Annotation										
NoSta-D-1.4-falko	10	6.034	0	0		Content Annotation										
NoSta-D-Anselm 2 2.710						Other Annotation										
NoSta-D-Kafka	10.388	0			All Annotation Diplomatic Text											
NoSta-D-TueBaDZ	2	10.832	0		Normalized Text											

Suchreferenz: https://korpling.org/annis3/?id=f7e2cc64-2800-4ba2-9830-f083fc1770af



 Suchen Sie einen Artikel direkt gefolgt von einem attribuierenden Adjektiv direkt gefolgt von einem Nomen!



- Suchen Sie einen Artikel direkt gefolgt von einem attribuierenden Adjektiv direkt gefolgt von einem Nomen!
 - Passendes Attribut wäre **pos.**



- Suchen Sie einen Artikel direkt gefolgt von einem attribuierenden Adjektiv direkt gefolgt von einem Nomen!
 - Passendes Attribut wäre **pos.**
 - Operator .
 - pos=/ART/ . pos=/ADJA/ . pos=/N./



🖃 🔿 🗮 About ANNIS		Report Pro	blem		Help us to make ANNIS better!	logged in as "laura" 🔒 Logout
					Help/Examples Q Query Result ×	
pos=/ART/				3	Base text 🗸	
pos=/ADJA/				Query	K (1 / 285) Displaying Results 1 - 10 of 2848	Result for: pos=/ART/ . pos=/ADJA/ . pos=/N./ 🔩
•				Juniaer	1 👩 < Path: RIDGES_Herbology_Version5.0-beta1 > AlchymistischePractic_1603_Libavius (dipl 190 - 202)	left context: 5 🗸 right context: 5 🗸
pos=/N./				******	publication zuwider fein / dafz der gemeine Deutfche nicht verftehe wie gelehrt fie	
					⊞ Text Layers	
					Lexical Annotation	
					Morphological Annotation	
					⊞ Syntactic Annotation	
Q Search More	•	History -	-		⊞ Graphical Annotation	
					Content Annotation	
2848 matches)	Other Annotation	
in 36 documents (showing	g 8/10) results)			All Annotation	
					Diplomatic Text	
Corpus List Search Option	ons				Normalized Text	
Visible: All			1	~ (C	2 👩 🛃 Path: RIDGES_Herbology_Version5.0-beta1 > AlchymistischePractic_1603_Libavius (dipl 541 - 553)	left context: 5 🗸 right context: 5 🗸
ridges					/ weit vnd enge / eine zimliche anzahl haben . Difz nennet man	
Name *	Texts	Tokens			⊞ Text Layers	
Ridges-SE-CompAnalyse	20	92,368	0	₽	Lexical Annotation	
ridges.herbology	14	63,734	0	F	Morphological Annotation	
RIDGES Herbology Version	22	122.698			Syntactic Annotation	
		122,000			H Graphical Annotation	
RIDGES_Herbology_Versior	29	154,266	0		Content Annotation	
RIDGES_Herbology_Versior	29	154,267	0			
RIDGES_Herbology_Versior	36	183,710	1	B		
Ridges Herbology Version	13	60.811	6		Normalized Text	
0 - 00- 1					3 A Path: RIDGES Herbology Version5.0-beta1 > AlchymistischePractic 1603 Libavius (dipl 594 - 607)	left context: 5 🗸 right context: 5 🗸
					ein Kurrel / vherlich mit einem land ren Halfz / werden Vind len renennet	
					E Tavt I work	
					Syntactic Apportation	
					Graphical Annotation	
					Other Annotation	

Suchreferenz: https://korpling.org/annis3/?id=e39f54cb-8a90-47d0-9a09-830c12022c32