

Carolin Odebrecht

Korpuslinguistik | Humboldt-Universität zu Berlin

# SUCHE NACH UND VISUALISIERUNG VON HISTORISCHEN MEHREBENENKORPORA MIT ANNIS

TraCES Projekt | Asien-Afrika-Institut der Universität Hamburg



# Überblick



## I. Suche nach historischen Korpora

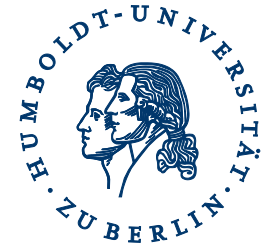
- ANNIS Interface
- ANNIS Query Language
  - Verkürzte Schreibweise, Spannenannotation
- Referenzen für Korpora, Suche und Treffer

## II. Visualisierung von historischen Korpora

- Textanzeige
- Multiple Segmentierungen
- Annotationsanzeigen
- Unterstützung von Schriftarten



# Überblick



## I. Suche nach historischen Korpora

- ANNIS Interface
- ANNIS Query Language
  - Verkürzte Schreibweise, Spannenannotation
- Referenzen für Korpora, Suche und Treffer

## II. Visualisierung von historischen Korpora

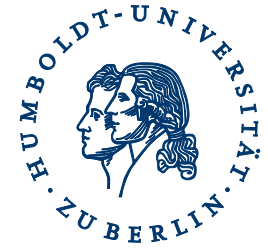
- Textanzeige
- Multiple Segmentierungen
- Annotationsanzeigen
- Unterstützung von Schriftarten



# SUCHE NACH HISTORISCHEN MEHREBENENKORPORA



# ANNIS



- Browser basiertes Such- und Visualisierungstool für Mehrebenenkorpora
  - Serverinstallation, lokale Installation
- generisches Datenmodell Salt (Zipser & Romary 2010)
  - Graph basiert (Krause & Zeldes 2014)
- generischen Anfragesprache ANNIS Query Language (AQL)
  - Zwei Schreibweisen: Klauselschreibweise und verkürzte Schreibweise
  - grundsätzlich gilt:
    - Gefunden werden kann nur das, was auch annotiert ist!**
  - kein Annotations- oder NLP-Tool



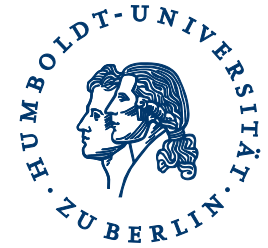
# ANNIS



- Öffentliche Instanzen, z.B.
  - Humboldt-Universität zu Berlin
    - Korpuslinguistik und Morphologie, Anke Lüdeling
    - <https://korpling.org/annis3/>
  - Georgetown University Washington
    - Computational Linguistics , Amir Zeldes
    - <https://corpling.uis.georgetown.edu/annis/scriptorium>
  - Universität Hamburg
    - Linguistik des Deutschen/Korpuslinguistik, Heike Zinsmeister
    - In Vorbereitung

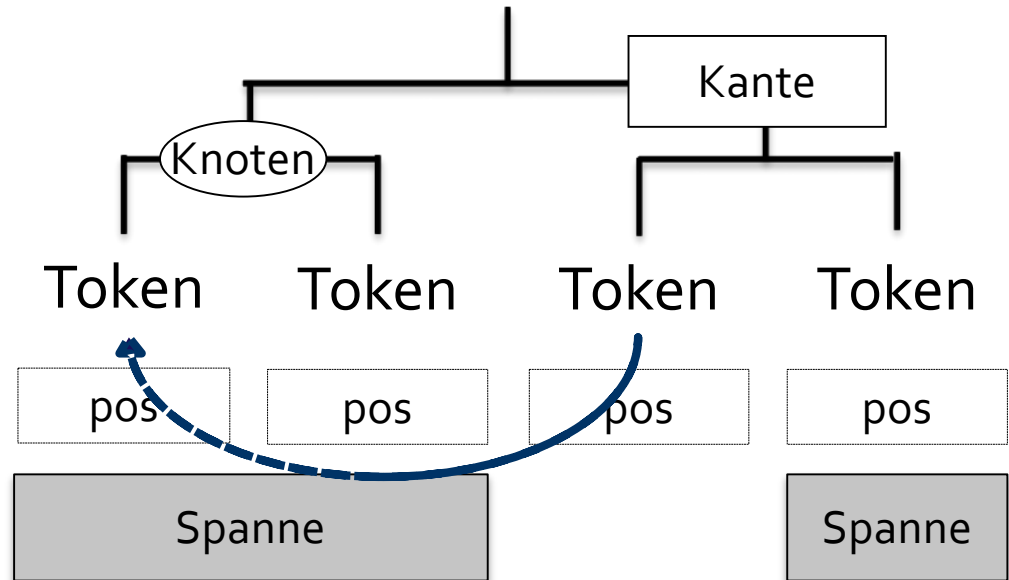


# Korpora Annotationen



- Suche und Visualisierung diverser Annotationskonzepte, z.B.
  - Token
  - Spannen
  - Bäume
  - Filterung nach Metadaten
- unabhängig von der Bedeutung der Annotationen (z.B. Tagset)

- Konverter Framework Pepper (Zipser & Romary 2010)
  - gemeinsames Datenmodell mit ANNIS → Salt
  - Unterstützung von u.a.
    - TEI XML,
    - MMAx,
    - EXMARaLDA,
    - ANNIS,
    - TIGER XML,
    - TCF,
    - PAULA







# Unser Zugang



- HU – Instanz
  - öffentlicher Zugang: <https://korpling.org/annis3/>
  - kein Login benötigt
- Suche mit zwei Korpora
  - Korpusreferenz RIDGES:  
[https://korpling.org/annis3/#\\_c=UklER0VTX0hlcmJvbG9neV9WZ\\_XJzaW9uNC4x](https://korpling.org/annis3/#_c=UklER0VTX0hlcmJvbG9neV9WZ_XJzaW9uNC4x)
  - Korpusreferenz Fürstinnen (Fürstinnenkorrespondenzkorpus):  
[https://korpling.org/annis3/#\\_c=RnVlcnN0aW5uZW5rb3JyZXNwb25kZW56MS4x](https://korpling.org/annis3/#_c=RnVlcnN0aW5uZW5rb3JyZXNwb25kZW56MS4x)
  - für alle Referenzlinks letzter Zugriff am 01.02.2016



# Erstes Beispielkorpus Fürstinnenkorrespondenz



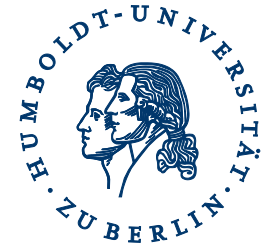
- Texte aus dem 17. und 19. Jahrhundert, ca. 260.000 Token
- Historisches Deutsch, Briefe adeliger Frauen
- Token- und Spannenannotationen:
  - U.a. annotiert mit Wortformen, Lemmata, Wortarten, Phrasen

|            |               |            |                |              |     |               |
|------------|---------------|------------|----------------|--------------|-----|---------------|
| grfunct    | prsimpl       | subj       |                |              |     | prsimpl       |
| lemma      | erfordern     | mein       | kindlich       | Schuldigkeit | ,   | wünschen      |
| mod_polite |               | n1-pn      | n1-fam         |              |     | x             |
| morph      | 3.Sg.Pres.Ind | Fem.Nom.Sg | Pos.Fem.Nom.Sg | Fem.Nom.Sg   |     | 1.Sg.Pres.Ind |
| norm       | erfordert     | meine      | kindliche      | Schuldigkeit | ,   | wünsche       |
| phras      | honor         |            |                |              |     | apprec        |
| pos        | VVFIN         | PPOSAT     | ADJA           | NN           | \$, | VVFIN         |
| tok        | Erfahrtert    | meine      | kindlige       | schuldigkeit | ,   | wünsche       |

Trefferreferenz: <https://korpling.org/annis3/?id=bd568035-4250-4e47-bb72-4e1dof4a482c>



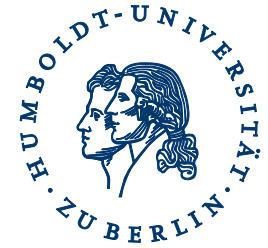
# Korpus Fürstinnenkorrespondenz



- Lühr, Rosemarie; Faßhauer, Vera; Prutscher, Daniela; Seidel, Henry; Fürstinnenkorrespondenz (Version 1.1), Universität Jena, DFG. <http://www.indogermanistik.uni-jena.de/Web/Projekte/Fuerstinnenkorr.htm>.  
<http://hdl.handle.net/11022/0000-0000-82A0-7>



# Zweites Beispielkorpus RIDGES



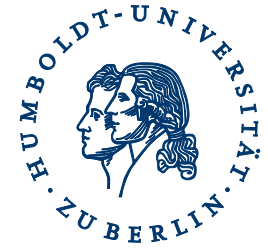
- Texte aus dem 15. und 19. Jahrhundert, ca. 150.000 Token
- Historisches Deutsch, Kräuterkundetexte
- Spannenannotationen
  - U.a. annotiert mit Wortformen, Wortarten, Zeilenumbrüchen, Autorenreferenz

|                   |         |        |        |         |        |        |        |        |        |                |          |
|-------------------|---------|--------|--------|---------|--------|--------|--------|--------|--------|----------------|----------|
| <b>author_ref</b> |         |        |        | pron1sg |        |        |        |        |        |                |          |
| <b>clean</b>      | .       | ES     | sindt  | mir     | lieber | Leser  | /      | zwey   | feine  | Alchymistische |          |
| <b>dipl</b>       | .       | ES     | findt  | mir     | lieber | Lefer  | /      | zwey   | feine  | Alchy          | mitfiche |
| <b>div2</b>       | div2    |        |        |         |        |        |        |        |        |                |          |
| <b>div2_type</b>  | preface |        |        |         |        |        |        |        |        |                |          |
| <b>head</b>       | head    |        |        |         |        |        |        |        |        |                |          |
| <b>hi</b>         |         | hi     |        |         |        |        |        |        |        |                |          |
| <b>hi_font</b>    | gothic  | gothic | gothic | gothic  | gothic | gothic | gothic | gothic | gothic | gothic         | gothic   |
| <b>lb</b>         | lb      | lb     |        |         |        |        |        |        |        |                | lb       |
| <b>lemma</b>      | .       | es     | sein   | ich     | lieb   | Leser  | /      | zwei   | fein   | alchymistisch  |          |
| <b>norm</b>       | .       | Es     | sind   | mir     | lieber | Leser  | /      | zwei   | feine  | Alchymistische |          |
| <b>p</b>          |         | p      |        |         |        |        |        |        |        |                |          |
| <b>pos</b>        | \$.     | PPER   | VAFIN  | PPER    | ADJD   | NN     | \$(    | CARD   | ADJA   | ADJA           |          |

Trefferreferenz: <https://korpling.org/annis3/?id=ab4cd8d6-a1be-4f2d-8372-e912ebfc9eb6>

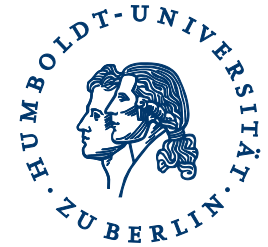


# Korpus RIDGES



- Lüdeling, Anke; Odebrecht, Carolin; Zeldes, Amir; RIDGES-Herbology (Version 4.1), Humboldt-Universität zu Berlin.  
<http://korpling.german.hu-berlin.de/ridges/>.  
<http://hdl.handle.net/11022/0000-0000-8253-F>
- vgl. Odebrecht et al. (eingereicht)

# Token



- 1) Als Token bezeichnet man häufig die **kleinste (technische) Einheit** in einem Korpus.
- 2) Ein Token entspricht oft (aber nicht immer) einem **orthographischen Wort** oder **Satzzeichen**.
- 3) Nach diesen Einheiten kann man in ANNIS **suchen**.

|            |               |            |                |              |            |                |
|------------|---------------|------------|----------------|--------------|------------|----------------|
| grfunct    | prsimpl       | subj       |                |              |            | prsimpl        |
| lemma      | erfordern     | mein       | kindlich       | Schuldigkeit | ,          | wünschen       |
| mod_polite |               | n1-pn      | n1-fam         |              |            | x              |
| morph      | 3.Sg.Pres.Ind | Fem.Nom.Sg | Pos.Fem.Nom.Sg | Fem.Nom.Sg   |            | 1.Sg.Pres.Ind  |
| norm       | erfordert     | meine      | kindliche      | Schuldigkeit | ,          | wünsche        |
| phras      | honor         |            |                |              |            | apprec         |
| pos        | VVEIN         | PROSAT     | ADIA           | NN           | ,          | VVEIN          |
| tok        | Erfohrtert    | meine      | kindlige       | schuldigkeit | ,          | wünsche        |
|            | <b>tok</b>    | <b>tok</b> | <b>tok</b>     | <b>tok</b>   | <b>tok</b> | <b>tok ...</b> |

# Spannenannotation



- 1) Ein oder mehreren Token wird ein Wert in einer Annotationsebene zugeordnet.

Wichtig: Nicht allen Token muss in jeder Ebene ein Wert zugeordnet werden!

|            |               |            |                |              |            |                |
|------------|---------------|------------|----------------|--------------|------------|----------------|
| grfunct    | prsimpl       | subj       |                |              |            | prsimpl        |
| lemma      | erfordern     | mein       | kindlich       | Schuldigkeit | ,          | wünschen       |
| mod_polite |               | n1-pn      | n1-fam         |              |            | x              |
| morph      | 3.Sg.Pres.Ind | Fem.Nom.Sg | Pos.Fem.Nom.Sg | Fem.Nom.Sg   |            | 1.Sg.Pres.Ind  |
| norm       | erfordert     | meine      | kindliche      | Schuldigkeit | ,          | wünsche        |
| phras      | honor         |            |                |              |            | apprec         |
| pos        | VVEIN         | PROSAT     | ADIA           | NN           | ,          | VVEIN          |
| tok        | Erfohrtert    | meine      | kindlige       | schuldigkeit | ,          | wünsche        |
|            | <b>tok</b>    | <b>tok</b> | <b>tok</b>     | <b>tok</b>   | <b>tok</b> | <b>tok ...</b> |



# ANNIS INTERFACE





# korpling.org/annis3



Navigation: About ANNIS | Report Problem | [Help us to make ANNIS better!](#) | not logged in | Login

Query Builder

Please enter AQL query

Search | More | History

Welcome to ANNIS! A tutorial is available on the right side.

Corpus List | Search Options

Visible: All

| Name                 | Texts | Tokens |  |  |
|----------------------|-------|--------|--|--|
| a5.hausa.news        | 4     | 2.017  |  |  |
| a5.hausa.umarnin.uwa | 47    | 10.194 |  |  |
| abraham.our.father   | 7     | 7.671  |  |  |
| AliBaba.6.all        | 1     | 202    |  |  |

Help/Examples

Tutorial

Example Queries

| Example Query                     | Description   | open corpus browser                  |
|-----------------------------------|---|--------------------------------------|
| <a href="#">Q /[/j]Jahr/</a>      | Search for the "jahr" with upper or lower-case "j" (regular expression)       | <a href="#">KAJUK</a>                |
| <a href="#">Q /[Ee]vangelika/</a> | Search for the "evangelika" with upper or lower-case "e" (regular expression) | <a href="#">a5.hausa.news</a>        |
| <a href="#">Q "."</a>             | search for the word "."   | <a href="#">a5.hausa.umarnin.uwa</a> |
| <a href="#">Q norm="νοϋτε"</a>    | search for the normalized word νοϋτε  | <a href="#">abraham.our.father</a>   |
| <a href="#">Q pos="NPRO"</a>      | search for proper names   | <a href="#">abraham.our.father</a>   |
| <a href="#">Q /[//]//</a>         | Search for the "/" with upper or lower-case "/" (regular expression)          | <a href="#">HSJ-Briefe</a>           |
| <a href="#">Q "zu"</a>            | search for the word "zu"  | <a href="#">HSJ-Maese_Nissima</a>    |
| <a href="#">Q /[Ff]un/</a>        | Search for the "fun" with upper or lower-case "f" (regular expression)        | <a href="#">HSJ-Maese_Nissima</a>    |
| <a href="#">Q "schwēster"</a>     | search for the word "schwēster"   | <a href="#">HSJ-Schevet_Jehude</a>   |
| <a href="#">Q /[Dd]as/</a>        | Search for the "das" with upper or lower-case "d" (regular expression)        | <a href="#">HSJ-Schevet_Jehude</a>   |
| <a href="#">Q "Aine"</a>          | search for the word "Aine"  | <a href="#">HSJ-Varia</a>            |
| <a href="#">Q /[Aa]inēr/</a>      | Search for the "ainer" with upper or lower-case "a" (regular expression)      | <a href="#">HSJ-Varia</a>            |
| <a href="#">Q  = "</a>            | Findet einen Junktor  | <a href="#">KAJUK</a>                |
| <a href="#">Q "in"</a>            | search for the word "in"  | <a href="#">KAJUK</a>                |



# Interface



ANNIS interface screenshot with annotations:

- Tutorial für erste Schritte**: Points to the [Tutorial](#) link in the Help/Examples menu.
- Beispielanfragen**: Points to the **Example Queries** section.
- Auswahl der Korpora für die Suche (mehrere mit Strg + Klick)**: Points to the **Visible:** dropdown menu in the Corpus List section.

**Corpus List**

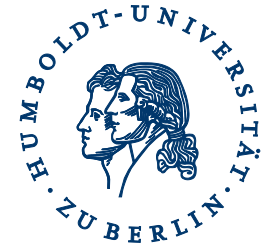
| Name                 | Texts | Tokens |   |   |
|----------------------|-------|--------|---|---|
| a5.hausa.news        | 4     | 2.017  | ⓘ | 📄 |
| a5.hausa.umarnin.uwa | 47    | 10.194 | ⓘ | 📄 |
| abraham.our.father   | 7     | 7.671  | ⓘ | 📄 |
| AliBaba.6.all        | 1     | 202    | ⓘ | 📄 |

**Example Queries**

| Example Query                   | Description  | open corpus browser                  |
|---------------------------------|--|--------------------------------------|
| <a href="#">/[/j]ahr/</a>       | Search for the lower-case "j"  |                                      |
| <a href="#">/[Ee]vangelika/</a> | Search for the or lower-case "e" (regular expression)                    | <a href="#">a5.hausa.news</a>        |
| <a href="#">".</a>              | search for the word "."  | <a href="#">a5.hausa.umarnin.uwa</a> |
| <a href="#">norm="νοϋτε"</a>    | search for the normalized word νοϋτε                                     | <a href="#">abraham.our.father</a>   |
| <a href="#">pos="NPRO"</a>      | search for proper names  | <a href="#">abraham.our.father</a>   |
| <a href="#">/[//]/</a>          | Search for the "/" with upper or lower-case "/" (regular expression)     | <a href="#">HSJ-Briefe</a>           |
| <a href="#">"zu"</a>            |  | <a href="#">jima</a>                 |
| <a href="#">/[Ff]un/</a>        |  | <a href="#">jima</a>                 |
| <a href="#">"schwēst"</a>       |  | <a href="#">Jude</a>                 |
| <a href="#">/[Dd]as/</a>        | Search for the "das" with upper or lower-case "d" (regular expression)   | <a href="#">HSJ-Schevet Jehude</a>   |
| <a href="#">"Aine"</a>          | search for the word "Aine"   | <a href="#">HSJ-Varia</a>            |
| <a href="#">/[Aa]iner/</a>      | Search for the "ainer" with upper or lower-case "a" (regular expression) | <a href="#">HSJ-Varia</a>            |
| <a href="#"> = "</a>            | Findet einen Junktor   | <a href="#">KAJUK</a>                |
| <a href="#">"in"</a>            | search for the word "in"   | <a href="#">KAJUK</a>                |



# Tutorial



Help us to make ANNIS better! not lo

Help/Examples

Tutorial

Choose topic ▾ [Print](#)

- ANNIS interface >
- ANNIS Query language >**

- Searching for Word Forms
- Searching for Annotations
- Searching using Regular Expressions
- Searching for Trees
- Searching for Pointing Relations
- Exporting Results
- Frequency Analysis
- Complete List of Operators

**Thema wählen**

ANNIS INTERFAC

## Using the ANNIS

The ANNIS interface i  
search form and the

### The Search Form

```
head_1_pos=/V.  
>dep[func="dobj"]
```

[Example Queries](#)



# Suchfenster



Fehler/Fragen  
an das  
ANNIS-Team

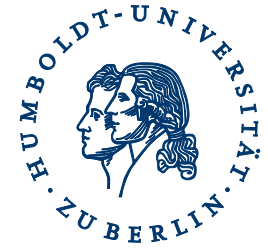
Shortcut Suche:  
Strg + Enter

Anzeige:

- a) Anzahl der Treffer
- b) Fehler in der  
Anfrage



# i-Button Metadaten



ANNIS interface showing a search query input field, a corpus list, and a table of example queries. A red arrow points to the 'i' icon in the corpus list, labeled '„i“ Zugriff auf Korpusmetadaten'.

Visible: All

| Name                 | Texts | Tokens | i | 📄 |
|----------------------|-------|--------|---|---|
| a5.hausa.news        | 4     | 2.017  | ⓘ | 📄 |
| a5.hausa.umarnin.uwa | 47    | 10.194 | ⓘ | 📄 |
| abraham.our.father   | 7     | 7.671  | ⓘ | 📄 |
| AliBaba.6.all        | 1     | 202    | ⓘ | 📄 |

| Example Query                     | Description   | open corpus browser                  |
|-----------------------------------|---|--------------------------------------|
| <a href="#">Q /[/j]Jahr/</a>      | Search for the "jahr" with upper or lower-case "j" (regular expression)       | <a href="#">KAJUK</a>                |
| <a href="#">Q /[Ee]vangelika/</a> | Search for the "evangelika" with upper or lower-case "e" (regular expression) | <a href="#">a5.hausa.news</a>        |
| <a href="#">Q "."</a>             | search for the word "."   | <a href="#">a5.hausa.umarnin.uwa</a> |
| <a href="#">Q norm="νοϋτε"</a>    | search for the normalized word νοϋτε  | <a href="#">abraham.our.father</a>   |
| <a href="#">Q pos="NPROPN"</a>    | search for proper names   | <a href="#">abraham.our.father</a>   |
| <a href="#">Q /[/]/</a>           | Search for the "/" with upper or lower-case "/" (regular expression)          | <a href="#">HSJ-Briefe</a>           |
| <a href="#">Q "zu"</a>            | search for the word "zu"  | <a href="#">HSJ-Maese_Nissima</a>    |
| <a href="#">Q /[/f]fun/</a>       | Search for the "fun" with upper or lower-case "f" (regular expression)        | <a href="#">HSJ-Maese_Nissima</a>    |
| <a href="#">Q "schwester"</a>     | Search for the word "schwester"   | <a href="#">HSJ-Schevet_Jehude</a>   |
| <a href="#">Q /[/Dd]das/</a>      | Search for the "das" with upper or lower-case "d" (regular expression)        | <a href="#">HSJ-Schevet_Jehude</a>   |
| <a href="#">Q "Aine"</a>          | Search for the word "Aine"  | <a href="#">HSJ-Varia</a>            |
| <a href="#">Q /[/Aa]ainer/</a>    | Search for the "ainer" with upper or lower-case "a" (regular expression)      | <a href="#">HSJ-Varia</a>            |
| <a href="#">Q j="j"</a>           | Findet einen Junktor  | <a href="#">KAJUK</a>                |
| <a href="#">Q "in"</a>            | search for the word "in"  | <a href="#">KAJUK</a>                |



# i-Button Korpusmetadaten



Corpus information for RIDGES\_Herbology\_Version4.1 (ID: 11636)

| Metadata  |   | Available annotations |  |
|---|---|-----------------------|--|
| Select corpus/document: RIDGES_Herbology_Version4.1 |   | Node Annotations      |  |
| Name  | Value   | Name                  | Example (click to use query)   |
| Homepage  | <a href="http://korpling.german.hu-berlin.de/ridges/index_en.html">http://korpling.german.hu-berlin.de/ridges/index_en.html</a>   | atLeast               | atLeast="1,000000"   |
| annotators  | Abed-Ali, Ilham; Andresen, Silke; Ast, Henriette; Belz, Malte; Christen, Doreen; Dayal, Mascha; Dittberner, Antonia; Döhn, Cora; Driemel, Imke; Efremova, Olja; Eichhorn, Gill-Maria; Esser, Judith; Gerlach, Annegret; Giesel, Linda; Kiraga, Sebastian; Kolbik, Ewa Anna; Kovács, Kornél; Krüger, Daisi; Lehmann, Anna-Maria; Lober, Maria; Lueders, Laura; Lüdeling, Anke, Maniscalco, Samuele; Metzig, Manuel; Meyer, Alexander; Müller, Sandra; Müller, Vinzent; Murphy, Andrew; Mursell, Johannes; Okuda, Akiko; Perlitz, Laura; Reinig, Katharina; Riesler, Ina; Rosin, Lena; Sachse, Franz-Josef; Sapronova, Anna; Sauer, Simon; Schmidt, Claudia; Sorokovska, Iryna; Springmann, Uwe; Stephan, Kristina; Tiemann, Juliane; Tóth, Anna; Tóth, Réka; Turtureanu, Alexander; Wekel, Juliana; Zuchewicz, Karolina et al. | atMost                | atMost="2,000000"  |
| projectDesc   | RIDGES Project (Register In Diachronic German Science), funded by the Google Digital Humanities Research Award  | attr_gen              | attr_gen="gpost"   |
| respStmt  | Compiled by Thomas Krause, Anke Lüdeling, Carolin Odebrecht, Amir Zeldes. The RIDGES team consists of Anke Lüdeling, Amir Zeldes, Carolin Odebrecht, Vivian Voigt and Laura Perlitz.  | author_ref            | author_ref="pron1sg"   |
| sponsor   | Google Inc.   | bemerkung             | bemerkung="grammatisch stellt es sich als verbalsubstantiv mit (idg.) ti-suffix zu got. siukan 'krank sein'. " |
| version   | 4.1   | brace                 | brace="brRight"  |
|   |   | brace_dir             | brace_dir="left"   |
|   |   | clean                 | clean="/"  |
|   |   | definition            | definition="fig"   |
|   |   | dipl                  | dipl="/"   |
|   |   | disease               | disease="di"   |
|   |   | div1                  | div1="div1"  |
|   |   | div1_tvne             | div1   |
|   |   | Edge Annotations      |  |
|   |   | Edge Types            |  |
|   |   | Meta Annotations      |  |

Link to corpus: [https://korpling.org/annis3/#\\_c=UkiER0VTX0hlcmljYbG9neV9WZXJzaW9uNC4x](https://korpling.org/annis3/#_c=UkiER0VTX0hlcmljYbG9neV9WZXJzaW9uNC4x)

**Korpusreferenzlink**



# i-Button

## Verfügbare Annotationen



attr\_gen="gpost"

Available annotations

| Name            | Example (click to use query)  | URL |
|-----------------|---|-----|
| atLeast         | atLeast="1,000000"  |     |
| atMost          | atMost="2,000000"   |     |
| <b>attr_gen</b> | <b>attr_gen="gpost"</b>   |     |
| author_ref      | author_ref="pron1sg"  |     |
| bemerkung       | bemerkung="grammatisch stellt es sich als verbalsubstantiv mit (idg.) ti-suffix zu got. siukan 'krank sein'." |     |
| brace           | brace="brRight"   |     |
| brace_dir       | brace_dir="left"  |     |
| clean           | clean="/"   |     |
| definition      | definition="fig"  |     |
| dipl            | dipl="/"  |     |
| disease         | disease="di"  |     |
| div1            | div1="div1"   |     |
| div1_tvne       | div1_tvne="book"  |     |

Link to corpus: [https://korpling.org/annis3/#\\_c=UkIER0VTX0hlcMjvbG9neV9WZjZjaW9uNC4x](https://korpling.org/annis3/#_c=UkIER0VTX0hlcMjvbG9neV9WZjZjaW9uNC4x)

Annotationsebenen und Beispiele für Werte. Doppelklick fügt das Beispiel in die Suchmaske ein.



# Trefferkonkordanz RIDGES



The screenshot displays the ANNIS interface with search results for 'RIDGES\_Herbology\_Version4.1'. The interface includes a top navigation bar with 'About ANNIS', 'Report Problem', and a 'Help us to make ANNIS better!' link. A search bar at the top right shows 'Query Result' and 'not logged in'. The main content area displays search results with document names, paths, and text snippets. Annotations are shown below each snippet, including Transcription, Lexical Annotation, Syntactical Annotation, Content Annotation, Graphical Annotation, and All. A corpus list on the left shows 1489 matches in 29 documents, with 'RIDGES\_Herbology\_Version4.1' highlighted. Red arrows point to specific elements: 'Suchreferenzlink' (search reference link), 'i-Button für Dokumentmetadaten' (i-button for document metadata), 'Dokumentname' (document name), 'Annotationsanzeigen' (annotation displays), 'Trefferreferenzlink' (hit reference link), and 'Treffer' (hit).

Suchreferenzlink

i-Button für Dokumentmetadaten

Dokumentname

Annotationsanzeigen

Trefferreferenzlink

Treffer

Trefferanzahl

| Name                        | Texts | Tokens  |  |  |
|-----------------------------|-------|---------|--|--|
| Fuerstinnenkorrespondenz    | 600   | 262.465 |  |  |
| RIDGES_Herbology_Version4.1 | 29    | 154.267 |  |  |





# i-Button Dokumentmetadaten



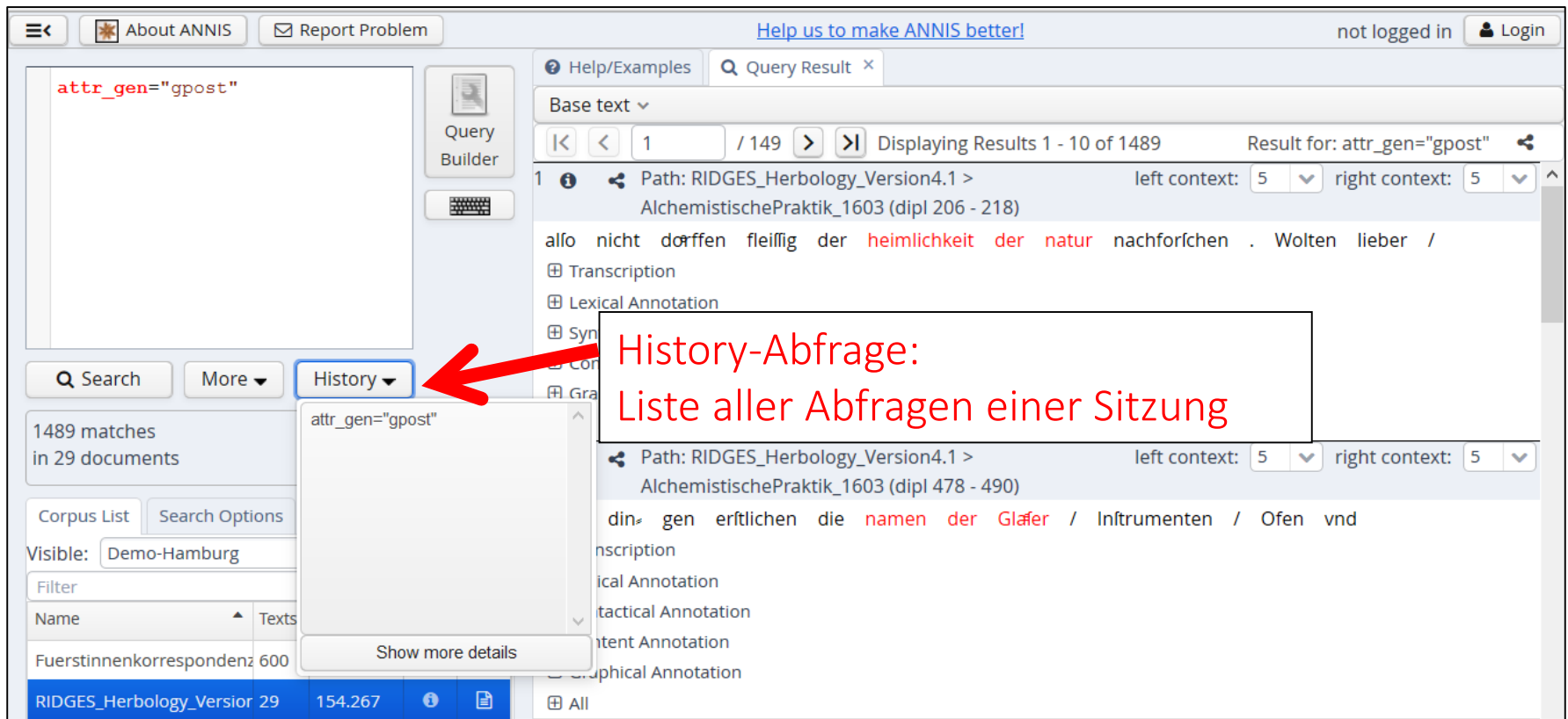
Info for salt:/RIDGES\_Herbology\_Version4.1/Alchemisti... + ×

Metadata

document: AlchemistischePraktik\_1603

| Name                 | Value  |
|----------------------|--|
| annis:doc            | AlchemistischePraktik_1603   |
| default_ns:author    | Andreas Libavius   |
| default_ns:bibl      | Libavius, Andreas (1603) Alchimistische Praktik. Frankfurt am Main. Johann Saur. 4-26. |
| default_ns:date      | 1603   |
| default_ns:pubPlace  | Frankfurt  |
| default_ns:publisher | Johann Saur  |
| default_ns:title     | Alchimistische Praktik   |
| default_ns:version   | 4.1  |

corpus: RIDGES\_Herbology\_Version4.1

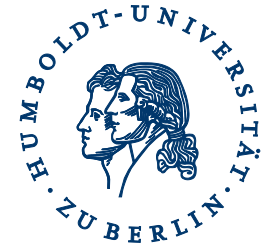


The screenshot shows the ANNIS web interface. At the top, there are navigation links like "About ANNIS" and "Report Problem", and a user status "not logged in" with a "Login" button. The main search area contains a text input field with the query `attr_gen="gpost"`. Below the input field are buttons for "Search", "More", and "History". A red arrow points to the "History" button, which has a dropdown menu open showing the same query `attr_gen="gpost"` and a "Show more details" button. The main content area displays search results for the query, showing document paths and snippets of text with highlighted search terms. A red text box with a white background is overlaid on the interface, containing the text: "History-Abfrage: Liste aller Abfragen einer Sitzung".



# Interface

## Anfrage Optionen



The screenshot shows the ANNIS search interface. At the top, there are navigation links for 'About ANNIS' and 'Report Problem'. The main search area contains the query `attr_gen="gpost"`. Below the search bar, there are buttons for 'Search', 'More', and 'History'. The results section shows '1489 matches in 29 documents'. The 'Search Options' tab is active, displaying several settings:

| Option           | Value     |
|------------------|-----------|
| Left Context     | 5         |
| Right Context    | 5         |
| Show context in  | dipl      |
| Results Per Page | 10        |
| Order            | Ascending |

Größe des Kontexts festlegen

Welche Transkriptionsebene soll in der Trefferliste angezeigt werden?

# Korpusliste



ANNIS interface showing a list of example queries and corpus data. The 'Visible' dropdown menu is highlighted with a red circle, showing 'All' selected and 'Demo-Hamburg' as an option.

Search Options

Visible: **All**

Filter: All

Name: **Demo-Hamburg**

| Example Query                      | Description   | open corpus browser                  |
|------------------------------------|---|--------------------------------------|
| <a href="#">Q /Ij]ahr/</a>         | Search for the "jahr" with upper or lower-case "j" (regular expression)       | <a href="#">KAJUK</a>                |
| <a href="#">Q /I[Ee]vangelika/</a> | Search for the "evangelika" with upper or lower-case "e" (regular expression) | <a href="#">a5.hausa.news</a>        |
| <a href="#">Q ".."</a>             | search for the word "."   | <a href="#">a5.hausa.umarnin.uwa</a> |
| <a href="#">Q norm="νοϋτε"</a>     | search for the normalized word νοϋτε  | <a href="#">abraham.our.father</a>   |
| <a href="#">Q pos="NPROPN"</a>     | search for proper names   | <a href="#">abraham.our.father</a>   |
| <a href="#">Q /I/I/I/</a>          | Search for the "/" with upper or lower-case "/" (regular expression)          | <a href="#">HSJ-Briefe</a>           |
| <a href="#">Q "zu"</a>             | search for the word "zu"  | <a href="#">HSJ-Maese Nissima</a>    |
| <a href="#">Q /I[F]un/</a>         | Search for the "fun" with upper or lower-case "f" (regular expression)        | <a href="#">HSJ-Maese Nissima</a>    |
| <a href="#">Q "schwēster"</a>      | search for the word "schwēster"   | <a href="#">HSJ-Schevet Jehude</a>   |
| <a href="#">Q /I[D]as/</a>         | Search for the "das" with upper or lower-case "d" (regular expression)        | <a href="#">HSJ-Schevet Jehude</a>   |
| <a href="#">Q "Aine"</a>           | search for the word "Aine"  | <a href="#">HSJ-Varia</a>            |
| <a href="#">Q /I[Aa]iner/</a>      | Search for the "ainer" with upper or lower-case "a" (regular expression)      | <a href="#">HSJ-Varia</a>            |
| <a href="#">Q ]="]"</a>            | Findet einen Junktor  | <a href="#">KAJUK</a>                |
| <a href="#">Q "in"</a>             | search for the word "in"  | <a href="#">KAJUK</a>                |

| Corpus Name          | Count | Size   | Info | Copy |
|----------------------|-------|--------|------|------|
| a5.hausa.news        | 4     | 2.617  | Info | Copy |
| a5.hausa.umarnin.uwa | 47    | 10.194 | Info | Copy |
| abraham.our.father   | 7     | 7.671  | Info | Copy |
| AliBaba.6.all        | 1     | 202    | Info | Copy |



# ANNIS QUERY LANGUAGE (AQL)



# Heute



- Demonstration von AQL mit zwei historischen Korpora
  - RIDGES und Fürstinnenkorrespondenz
  - Spannenannotation
  - Relationen zwischen Annotationen
  - Metadatenfilterung



# AQL



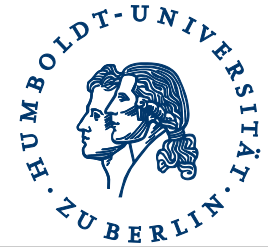
## ANNIS Query Language

- Prinzip I
    - Attribut-Wert-Paare
  - Prinzip II
    - Relationen
- Gilt für alle Annotationsarten und Korpora in ANNIS!



# Prinzip I

## Attribut-Wert-Paar



- 1) Voraussetzung ist das **Vorhandensein einer Ebene namens „dipl“**. (Metadaten!)
- 2) Erwartetes Ergebnis ist es, **exakt alle Vorkommen** dieser Zeichenkette in „dipl“ im ausgewählten Korpus zu finden.

dipl= /kraut/

**Attribut**

**Wert**

(Layer, Tier, Ebene ...)

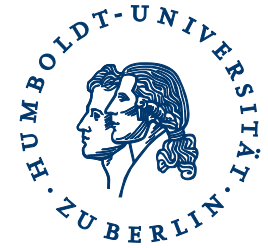
(Wort, Lemma, Satz, Wortart ...)

|              |        |      |     |          |              |              |   |      |    |        |    |
|--------------|--------|------|-----|----------|--------------|--------------|---|------|----|--------|----|
| <b>dipl</b>  | wider  | auff | ein | frisches | zerstoffenes | <b>kraut</b> | / | thue | es | wider  | in |
| <b>clean</b> | wider  | auff | ein | frisches | zerstossenes | kraut        | / | thue | es | wider  | in |
| <b>norm</b>  | wieder | auf  | ein | frisches | zerstoßenes  | Kraut        | / | tue  | es | wieder | in |



# Attribut-Wert-Paar

## Beispiel RIDGES



- Suche nach einer historischen Wortform von *Kräuter*
  - dipl
    - Annotationsebene mit Transkriptionen historischer Wortformen
  - eine historische Form von *Kräuter*
    - Kräut<sup>e</sup>r

dipl=/Kräut<sup>e</sup>r/

Suchreferenz: <https://korpling.org/annis3/?id=d8421ff7-cd56-4e01-a405-15b7cbe41e77>

---

# Attribut-Wert-Paar

## Beispiel RIDGES



- Suche nach einer historischen Wortform von *und*
  - dipl
    - Annotationsebene mit Transkriptionen historischer Wortformen
  - eine historische Form von *und*
    - vnnd

dipl=/vnnd/

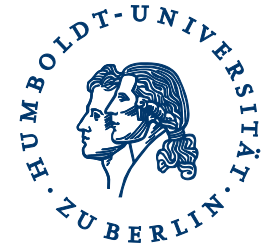
Suchreferenz: <https://korpling.org/annis3/?id=9e40ee93-748a-4fee-9fa7-01106ab453c3>

---



# EXKURS: REFERENZIERUNG IN ANNIS

# Korpusreferenz



- Aufrufen von ANNIS + ein Korpus ist ausgewählt
  - RIDGES in ANNIS
  - [https://korpling.org/annis3/#\\_c=UklER0VTX0hIcmJvbG9neV9WZXJzaW9uNC4x](https://korpling.org/annis3/#_c=UklER0VTX0hIcmJvbG9neV9WZXJzaW9uNC4x)
  - Link
    - Korpusmetadaten

| Corpus Information for RIDGES_Herbology_Verslon4.1 (ID: 11636) |   | Available annotations |   |
|--|---|-----------------------|---|
| Metadata   |   | Node Annotations      |   |
| Select corpus/document:  | RIDGES_Herbology_Verslon4.1   | Name                  | Example (click to use query)  |
| Name   | Value   | atLeast               | atLeast="1,000000"  |
| Homepage   | <a href="http://korpling.german.hu-berlin.de/ridges/index_en.html">http://korpling.german.hu-berlin.de/ridges/index_en.html</a>   | atMost                | atMost="2,000000"   |
| annotators   | Abed-All, Ilham; Andresen, Silke; Ast, Henriette; Belz, Malte; Christen, Doreen; Dayal, Mascha; Dittbener, Antonia; Dohn, Cora; Driemel, Imke; Efremova, Olga; Eichhorn, Gill-Maria; Esser, Judith; Gerlach, Annegret; Giesel, Linda; Kiraga, Sebastian; Kolbik, Ewa Anna; Kovacs, Kornel; Krüger, Daisy; Lehmann, Anna-Maria; Lober, Maria; Lueders, Laura; Ludeling, Anke; Mantisalco, Samuele; Metzig, Manuel; Meyer, Alexander; Müller, Sandra; Müller, Vinzent; Murphy, Andrew; Mursell, Johannes; Okuda, Akiko; Perlitz, Laura; Reinig, Katharina; Riesler, Ina; Rosin, Lena; Sachse, Franz-Josef; Sapronova, Anna; Sauer, Simon; Schmidt, Claudia; Sorokovska, Iryna; Springmann, Uwe; Stephan, Kristina; Tiemann, Juliane; Tóth, Anna; Tóth, Réka; Turtureanu, Alexander; Weikel, Juliana; Zuchewicz, Karolina et al. | attr_gen              | attr_gen="gpost"  |
| projectDesc  | RIDGES Project (Register in Diachronic German Science), funded by the Google Digital Humanities Research Award. Compiled by Thomas Krause, Anke Ludeling, Carolin Odebrecht, Amir Zeldes. The RIDGES team consists of Anke Ludeling, Amir Zeldes, Carolin Odebrecht, Vivian Voigt and Laura Perlitz.  | author_ref            | author_ref="pron1sg"  |
| respStmnt  | Google Inc.   | bemerkung             | bemerkung="grammatisch stellt es sich als verbalsubstantiv mit (ldg.) ti-suffix zu got. siukan 'krank sein'."   |
| version  | 4.1   | brace                 | brace="brRight"   |
|  |   | brace_dir             | brace_dir="left"  |
|  |   | clean                 | clean="/"   |
|  |   | definition            | definition="fig"  |
|  |   | dipl                  | dipl="/"  |
|  |   | disease               | disease="dl"  |
|  |   | div1                  | div1="div1"   |
|  |   | div1_tvne             | div1_tvne="hbnk"  |
|  |   | Edge Annotations      |   |
|  |   | Edge Types            |   |
|  |   | Meta Annotations      |   |
|  |   |                       | Link to corpus: <a href="https://korpling.org/annis3/#_c=UklER0VTX0hIcmJvbG9neV9WZXJzaW9uNC4x">https://korpling.org/annis3/#_c=UklER0VTX0hIcmJvbG9neV9WZXJzaW9uNC4x</a> |

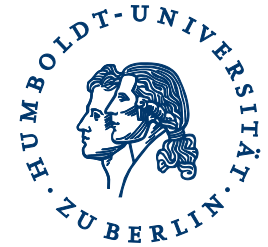
# Suchreferenz



- Aufrufen von ANNIS + ein Korpus ist ausgewählt + eine Anfrage ist gestellt
  - RIDGES in ANNIS + Suche `dipl=/vnnd/`
  - <https://korpling.org/annis3/?id=3e139941-0998-47fa-a59a-fb3f39cb15b3>
  - Link-Symbol
    - Trefferansicht


The screenshot shows the ANNIS search interface. At the top, there is a search bar with the query `dipl=/vnnd/` and a search button. Below the search bar, the results are displayed. The first result is highlighted and shows the following text: `allerley handt griff zuverftehen / vnnd gute Artzneyen zube reiten /`. Below the text, there are several expandable annotations: Transcription, Lexical Annotation, Syntactical Annotation, Content Annotation, Graphical Annotation, and All. The second result is also visible, showing the text: `wie gelehrt fie fein / vnnd alfo nicht dörffen fleißig der`. The interface also includes navigation controls and a status bar at the bottom.

# Trefferreferenz




- Aufrufen von ANNIS + ein Korpus ist ausgewählt + eine Anfrage ist gestellt
  - RIDGES in ANNIS + Suche `dipl=/vnnd/` + 1.Treffer
  - <https://korpling.org/annis3/?id=d136b502-6f44-4779-a323-4797987b4af9>
  - Link-Symbol
    - Treffer

Base text ▾

1  Path: RIDGES\_Herbiology\_Version4.1 > left context: 5 right context: 5  
AlchemistischePraktik\_1603 (dipl 34 - 44)

allerley handt griff zuverftehen / vnnd gute Artzneyen zube reiten /

⊕ Transcription  
⊕ Lexical Annotation  
⊕ Syntactical Annotation  
⊕ Content Annotation  
⊕ Graphical Annotation  
⊕ All

2  Path: RIDGES\_Herbiology\_Version4.1 > left context: 5 right context: 5  
AlchemistischePraktik\_1603 (dipl 200 - 210)

wie gelehrt fie fein / vnnd alfo nicht dörffen fleißig der

⊕ Transcription  
⊕ Lexical Annotation  
⊕ Syntactical Annotation  
⊕ Content Annotation  
⊕ Graphical Annotation  
⊕ All

# Trefferreferenz Einstellung



Match reference link + x

Share your match: 1. Choose the visualization to share. 2. Copy the generated link or code. 3. Share this link with your peers or include the code in your website.

Select visualization

|                        |
|------------------------|
| Display Name           |
| <b>kwic</b>            |
| Transcription          |
| Lexical Annotation     |
| Syntactical Annotation |
| Content Annotation     |
| Graphical Annotation   |
| All                    |

Link for publications

<https://korpling.org/annis3/?id=52b4dee4-f15d-4d61-a44a-5e848f9797cd>

Code for embedding visualization into web page

```
<iframe height="300px" width="100%" src="https://korpling.org/annis3/?id=52b4dee4-f15d-4d61-a44a-5e848f9797cd"></iframe>
```

Preview

[Show in ANNIS search interface](#)

allerley handt griff zuverftehen / vnnd gute Artzneyen zube

< >

Close

Links für  
Publikationen und  
Homepages

Auswahl der  
Annotationsanzeige

Vorschau

# Trefferreferenz Ausführen



[Link zur Suchanfrage](#)

[\\* Show in ANNIS search interface](#) 

|                  |          |            |             |         |        |        |           |            |         |        |
|------------------|----------|------------|-------------|---------|--------|--------|-----------|------------|---------|--------|
| <b>clean</b>     | allerley | handtgriff | zuverstehen | /       | vnnd   | gute   | Artzneyen | zube-      | reiten  | /      |
| <b>dipl</b>      | allerley | handtgriff | zuverstehen | /       | vnnd   | gute   | Artzneyen | zube-      | reiten  | /      |
| <b>div2</b>      | div2     |            |             |         |        |        |           |            |         |        |
| <b>div2_type</b> | preface  |            |             |         |        |        |           |            |         |        |
| <b>hi_font</b>   | gothic   | gothic     | gothic      | gothic  | gothic | gothic | gothic    | gothic     | gothic  | gothic |
| <b>komp</b>      |          | k          |             |         |        |        |           |            |         |        |
| <b>komp_orth</b> |          | lb2        |             |         |        |        |           |            |         |        |
| <b>lb</b>        | lb       | lb         |             |         |        |        |           |            |         | lb     |
| <b>lemma</b>     | allerlei | Handgriff  | <unknown>   | /       | und    | gut    | Arznei    | zubereiten | /       |        |
| <b>norm</b>      | allerlei | Handgriff  | zuverstehen | /       | und    | guten  | Arzneien  | zubereiten | /       |        |
| <b>p</b>         | p        |            |             |         |        |        |           |            |         |        |
| <b>pos</b>       | PIAT     | NN         | ADJA        | \$(     | KON    | ADJA   | NN        | VVINF      | \$(     |        |
| <b>pos_klein</b> | ART      | N          | ADJ         | ZEICHEN | KOORD  | ADJ    | N         | VINF       | ZEICHEN |        |
| <b>prot</b>      |          | prot1      |             |         |        |        |           |            |         |        |

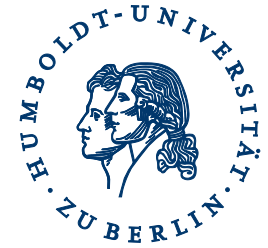




ZURÜCK ZU AQL

# Attribut-Wert-Paar

## Beispiel Fürstinnen



- Suche nach einer historischen Wortform von *und*
  - tok
    - Annotationsebene mit Transkriptionen historischer Wortformen
  - eine historische Form von *und*
    - *vnnd*

tok=/vnnd/

Suchreferenz: <https://korpling.org/annis3/?id=fff25c93-1a3c-4997-99f4-e67f23901210>

---

# Attribut-Wert-Paar

## Beispiel RIDGES



- Suche nach allen historischen Wortformen von *und*
  - norm
    - orthographische Normalisierung
  - historische Formen von *und*
    - *vnnd, vnd, vn, und ...*

norm=/und/

Suchreferenz: <https://korpling.org/annis3/?id=80f50383-67fc-415d-8d2a-62dcbd4bd0d9>

---

# Attribut-Wert-Paar

## Beispiel Fürstinnen



- Suche nach allen historischen Wortformen von *und*
  - norm
    - orthographische Normalisierung
  - historische Formen von *und*
    - *vnnd, vnd, vn, und ...*

norm=/und/

Suchreferenz: <https://korpling.org/annis3/?id=do1ffe16-ae61-49d6-8e0b-00359fa5e230>

---

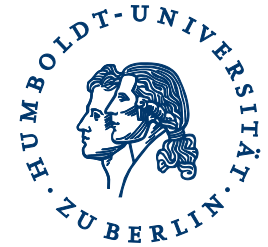
# Operatoren für die Mustersuche



- . 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
- (a | b) a oder b (auch: [ab])

# Operatoren

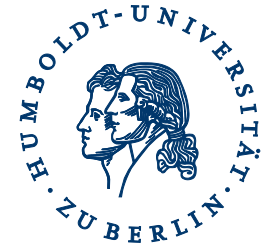
## Beispiel RIDGES



- Welche Ergebnisse erwarten Sie für folgende Anfragen, sogenannte Mustersuchen?
  - norm=/g.b./
  - <https://korpling.org/annis3/?id=e505b27c-b8ae-4113-9542-9158e9ef5704>
  - *gebe, gibt* (für RIDGES, theoretisch noch andere möglich)
  
  - dipl=/r(a|o)t/
  - <https://korpling.org/annis3/?id=b1e0d8c5-8a48-4152-b103-ef5ed9247af3>
  - *rot, rat* (für RIDGES)
  
  - dipl=/meint?/
  - <https://korpling.org/annis3/?id=7067989c-db93-41d4-a4c2-d6e93021cf91>
  - *mein, meint* (für RIDGES)

# Prinzip II

## Relationen RIDGES



head=/head/      AW-Paar1

o

Relation

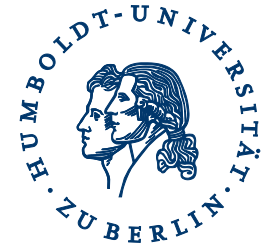
dipl=/Wermut/      AW-Paar2

- 1) Es gibt **mehrere Arten** von Annotationen!
- 2) Wissen, wie diese Annotationen in **Beziehung** zu einander stehen können!
- 3) **Spanne** (head) überlappt **dipl-Einheit** (dipl)!

|      |    |       |     |                |   |            |         |   |        |   |
|------|----|-------|-----|----------------|---|------------|---------|---|--------|---|
| dipl | an | ftatt | des | Zittwer-Samens | . | Abfinthium | Vulgare | , | Wermut | . |
| lb   | lb |       |     |                |   | lb         |         |   |        |   |
| head |    |       |     |                |   | head       |         |   |        |   |

Trefferreferenz: <https://korpling.org/annis3/?id=24f96036-d35a-42ea-8396-427e5022b31f>

# Syntax-Highlighting



- pro AW-Paar eine Farbe:
  - Farbe **Rot** für alle Werte des Attributs **head**
  - Farbe **Lila** für alle Werte des Attributs **dipl**
- Treffer in genau diesen Farben:
  - **rot** für die Überschrift
  - **lila** für die dipl-Einheit „Wermut“

```
head=/head/  
  o  
dipl=/Wermut/
```

Abfinthium Vulgare , Wermut .

|             |    |       |     |                |   |            |         |   |        |   |
|-------------|----|-------|-----|----------------|---|------------|---------|---|--------|---|
| <b>dipl</b> | an | ftatt | des | Zittwer-Samens | . | Abfinthium | Vulgare | , | Wermut | . |
| <b>lb</b>   | lb |       |     |                |   | lb         |         |   |        |   |
| <b>head</b> |    |       |     |                |   | head       |         |   |        |   |



# wichtige Relationen zwischen AW-Paaren

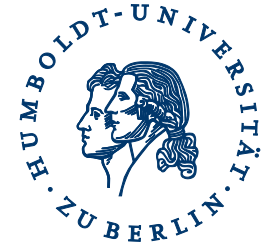


- . Direkte Präzedenz
- .\* Indirekte Präzedenz
- =\_= Identische Abdeckung
- \_i\_ Inklusion
- \_o\_ Überlappung
- \_l\_ linksseitige Überlappung
- \_r\_ rechtsseitige Überlappung

➤ Liste aller Operatoren im Tutorial!

# Relationen

## Beispiel RIDGES



- Suche nach Phrasen in Form von Wortartenabfolgen
  - pos
    - Annotation der Wortarten
    - Artikel Adjektiv Nomen = NP
  - Relation: direkte Präzedenz

pos=/ART/ . pos=/ADJA/ . pos=/NN/

Suchreferenz: <https://korpling.org/annis3/?id=93335489-gf6e-4491-ab47-c939cbd3ac65>

---

# Relationen

## Beispiel RIDGES



ANNIS interface showing search results for the query `pos=/ART/ . pos=/ADJA/ . pos=/NN/`.

Search results summary: 2331 matches in 29 documents.

Visible: Demo-Hamburg

| Name                        | Texts   | Tokens  |   |   |
|-----------------------------|---------|---------|---|---|
| Fuerstinnenkorrespondenz    | 600     | 262.465 | i | d |
| RIDGES_Herbology_Versior 29 | 154.267 |         | i | d |

Query Builder

Help/Examples | Query Result x

Base text v

1 Path: RIDGES\_Herbology\_Version4.1 > left context: 5 right context: 5  
AlchemistischePraktik\_1603 (dipl 190 - 202)  
publication zuwider fein / daz der gemeine Deutsche nicht verthehe wie gelehrt fie  
⊕ Transcription  
⊕ Lexical Annotation  
⊕ Syntactical Annotation  
⊕ Content Annotation  
⊕ Graphical Annotation  
⊕ All

2 Path: RIDGES\_Herbology\_Version4.1 > left context: 5 right context: 5  
AlchemistischePraktik\_1603 (dipl 541 - 553)  
/ weit vnd enge / eine zimliche anzahl haben . Difiz nennet man  
⊕ Transcription  
⊕ Lexical Annotation  
⊕ Syntactical Annotation  
⊕ Content Annotation  
⊕ Graphical Annotation  
⊕ All

3 Path: RIDGES\_Herbology\_Version4.1 > left context: 5 right context: 5  
AlchemistischePraktik\_1603 (dipl 595 - 608)  
Kugel / vber sich mit einem lan gen Halfz / werden Vio len genennet  
⊕ Transcription

# Relationen

## Beispiel Fürstinnen



- Suche nach Phrasen in Form von Wortartenabfolgen
  - pos
    - Annotation der Wortarten
    - Artikel Adjektiv Nomen = NP
  - Relation: direkte Präzedenz

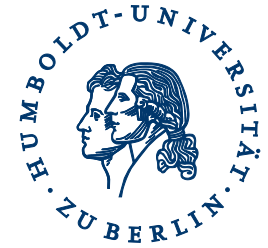
pos=/ART/ . pos=/ADJA/ . pos=/NN/

Suchreferenz: <https://korpling.org/annis3/?id=35040bf5-0b04-4bca-8f87-42504093fdf9>

---

# Relationen

## Beispiel Fürstinnen



[About ANNIS](#)
[Report Problem](#)
[Help us to make ANNIS better!](#)
not logged in [Login](#)

`pos=/ART/ . pos=/ADJA/ .`  
`pos=/NN/`

2599 matches in 537 documents

Visible:

| Name                     | Texts | Tokens  |                   |                   |
|--------------------------|-------|---------|-------------------|-------------------|
| Fuerstinnenkorrespondenz | 600   | 262.465 | <a href="#">i</a> | <a href="#">d</a> |
| RIDGES_Herbology_Versior | 29    | 154.267 | <a href="#">i</a> | <a href="#">d</a> |

[Help/Examples](#)
[Query Result](#)

Base text ▾

/ 260
 Displaying Results 1 - 10 of 2599

**1** Path: Fuerstinnenkorrespondenz1.1 > AD > AD\_JE2\_1677\_08\_14 (tokens 223 - 235)
 left context: 5 right context: 5

, so hätte Ich wohl **einen guten Peltz** , fon nöhten , AIß

⊕ general (grid)

**2** Path: Fuerstinnenkorrespondenz1.1 > AD > AD\_JE2\_1677\_08\_14 (tokens 436 - 448)
 left context: 5 right context: 5

mihz zu troste , und **den gantzen lande** , zu grossen Auffnehmmen ,

⊕ general (grid)

**3** Path: Fuerstinnenkorrespondenz1.1 > AD > AD\_JE2\_1677\_11\_07 (tokens 73 - 85)
 left context: 5 right context: 5

Grund meiner Sellen , das **der Allmächtige Gott** , Ew Gnd, ihn gleigen

⊕ general (grid)

**4** Path: Fuerstinnenkorrespondenz1.1 > AD > AD\_JE2\_1677\_11\_07 (tokens 82 - 94)
 left context: 5 right context: 5

Ew Gnd, ihn gleigen Auch **die Libe Frau** Mutter Gd mit seyner himmel

⊕ general (grid)

**5** Path: Fuerstinnenkorrespondenz1.1 > AD > AD\_JE2\_1677\_11\_07 (tokens 341 - 353)
 left context: 5 right context: 5

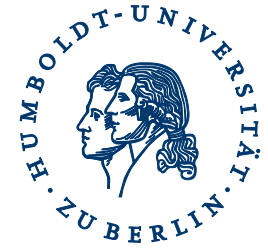
Princesin von Bernburg Id, sagen **Einen demütgen Gruß** , beÿ Ihr Gnd der

⊕ general (grid)

**6** Path: Fuerstinnenkorrespondenz1.1 > AD > AD\_JE2\_1677\_11\_07 (tokens 265 - 277)
 left context: 5 right context: 5

# Relationen

## Beispiel RIDGES



- Suche nach allen historischen Varianten von *und*
  - dipl
    - historische Formen von *und*
    - *vnnd, vnd, vn, und ...*
  - norm
    - orthographische Normalisierung
  - Relation: Überlappung

dipl

\_o\_

norm=/und/

Suchreferenz: <https://korpling.org/annis3/?id=1bf58b18-daa3-44bo-86ae-4c9db5a79da4>

---

# Relationen

## Frequenzanalyse



- Suche mit Frequenzanalyse kombinieren
  - Vorkommen der einzelnen Schreibvarianten von *und*

Anfrage bereits gestellt

dipl  
\_o\_  
norm=/und/

5117 matches in 29 documents

| Name                     | Texts | Tokens  |
|--------------------------|-------|---------|
| Fuerstinnenkorrespondenz | 600   | 262.465 |
| RIDGES_Herbology_Versior | 29    | 154.267 |

Export  
Frequency Analysis

Path: RIDGES\_Herbology\_Version4.1 >  
AlchemistischePraktik\_1603 (dipl 41 - 51)

Path: RIDGES\_Herbology\_Version4.1 >  
AlchemistischePraktik\_1603 (dipl 114 - 124)

# Frequenzanalyse Einstellung



ANNIS interface showing search results and frequency analysis settings.

Query: `dipl`  
Norm: `norm=/und/`

selected corpora: RIDGES\_Herbology\_Version4.1  
query to analyze: `dipl_o_norm=/und/`

|   | Node number/n | Selected annotation of node | Comment  |
|---|---------------|-----------------------------|--|
| 1 | 1             | dipl                        | automatically created from <code>dipl</code>       |
| 2 | 2             | norm                        | automatically created from <code>norm=/und/</code> |

5117 matches in 29 documents

Visible: Demo-Hamburg

| Name                        | Texts | Tokens  |   |   |
|-----------------------------|-------|---------|---|---|
| Fuerstinnenkorrespondenz    | 600   | 262.465 | i | d |
| RIDGES_Herbology_Versior 29 | 29    | 154.267 | i | d |

Start der Frequency Analysis

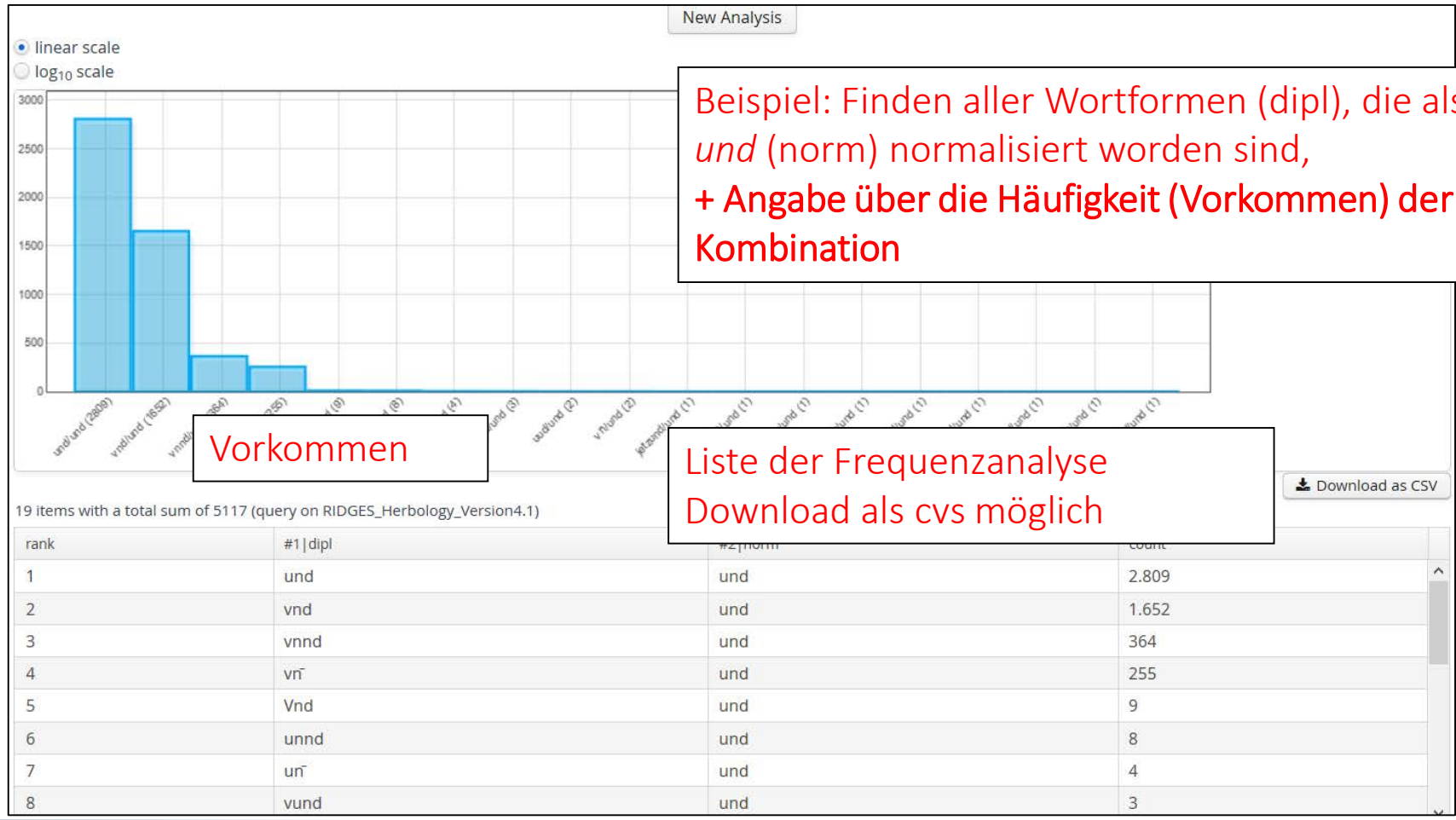
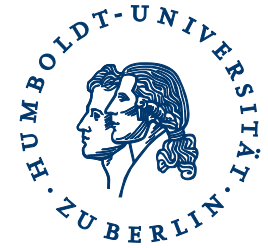
Perform frequency analysis

Beispiel: Finden aller Wortformen (`dipl`), die als `und` (`norm`) normalisiert worden sind

Start der Frequency Analysis



# Frequenzanalyse Ergebnis



# Relationen

## Beispiel Fürstinnen



- Suche nach allen historischen Varianten von *und*
  - tok
    - historische Formen von *und*
    - *vnnd, vnd, vn, und ...*
  - norm
    - orthographische Normalisierung
  - Relation: Überlappung

tok

\_o\_

norm=/und/

Suchreferenz: <https://korpling.org/annis3/?id=021532a2-0e6c-418f-a725-b6976c340eee>

---

# Filterung nach Metadaten

## RIDGES



- Suche nach allen historischen Varianten von *und* im 16. Jh.
  - dipl
    - historische Formen von *und*
    - *vnnd, vnd, vn, und ...*
  - norm
    - orthographische Normalisierung
  - Relation: Überlappung, Datumsangabe

dipl

\_o\_

norm=/und/

& meta::date=/15.\*/

Suchreferenz: <https://korpling.org/annis3/?id=483963d5-f007-4825-ged5-e3750acoda1e>

---

# Filterung nach Metadaten Fürstinnen

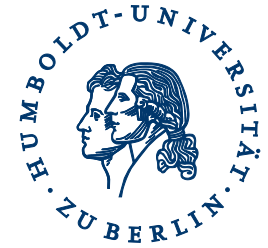


- Suche nach allen historischen Varianten von *und* im 16. Jh.
  - tok
    - historische Formen von *und*
    - *vnnd, vnd, vn, und ...*
  - norm
    - orthographische Normalisierung
  - Relation: Überlappung, Datumsangabe

```
tok
_o_
norm=/und/
& meta::date=/15.*/
```

Suchreferenz: <https://korpling.org/annis3/?id=787c6d37-gfc2-4e0a-a140-aa2386813obb>

---



# VISUALISIERUNG VON HISTORISCHEN MEHREBENENKORPORA



# Überblick



## I. Suche nach historischen Korpora

- ANNIS Interface
- ANNIS Query Language
  - Verkürzte Schreibweise, Spannenannotation
- Referenzen für Korpora, Suche und Treffer

## II. Visualisierung von historischen Korpora

- Textanzeige
- Multiple Segmentierungen
- Annotationsanzeigen
- Unterstützung von Schriftarten



# TEXTANZEIGE

# Kontext

- Ergebnis einer Suche sind Treffer mit
  - Annotationsanzeige, key word in context (kwic)
  - mit Kontexteinstellungen
- Kontext über den Treffer hinaus
  - Dokumentenbrowser (unabhängig von der Suche)
  - Annotationsanzeige für Textebenen (für jeden Treffer möglich)
  - Einbindung anderer Annotationen
  - Dokumentation unter [http://corpus-tools.org/annis/resources/ANNIS HTML Vis Guide 1.1.0.pdf](http://corpus-tools.org/annis/resources/ANNIS_HTML_Vis_Guide_1.1.0.pdf)
  - Beide Textanzeigen basieren auf den im Korpus vorhandenen Annotationen!



# Dokumentenbrowser

## Beispiel RIDGES



not logged in Login

Help/Examples Frequency Analysis Query Result Fuerstinnenkorrespondenz1.1 RIDGES

Filter documents by name

| corpus path  | visualizer   | info |
|--|--|------|
| RIDGES_Herbology_Version4.1 > AlchemistischePraktik_1603                 | <a href="#">diplomatic transcription</a><br><a href="#">normalized transcription</a> |      |
| RIDGES_Herbology_Version4.1 > ArzneiderKreutter_1532                     | <a href="#">diplomatic transcription</a><br><a href="#">normalized transcription</a> |      |
| RIDGES_Herbology_Version4.1 > ContrafaytKreuterbuch_1532                 | <a href="#">diplomatic transcription</a><br><a href="#">normalized transcription</a> |      |
| RIDGES_Herbology_Version4.1 > DerSchweizerischeBotanicus_1687            | <a href="#">diplomatic transcription</a><br><a href="#">normalized transcription</a> |      |
| RIDGES_Herbology_Version4.1 > DeutschePflanzennamen_1870                 | <a href="#">diplomatic transcription</a><br><a href="#">normalized transcription</a> |      |
| 828 RIDGES_Herbology_Version4.1 > DieEigenschaftenAllerHeilpflanzen_1828 | <a href="#">diplomatic transcription</a><br><a href="#">normalized transcription</a> |      |
| 764 RIDGES_Herbology_Version4.1 > EinleitungZuDerKraeuterkenntnis_1764   | <a href="#">diplomatic transcription</a><br><a href="#">normalized transcription</a> |      |
| RIDGES_Herbology_Version4.1 > Flora-6_1821                               | <a href="#">diplomatic transcription</a><br><a href="#">normalized transcription</a> |      |

Query Builder

Please enter AQL query

Search More History

Empty query

Corpus List Search Options

Visible: Demo-Hamburg

| Name                        | Texts   | Tokens  |  |  |
|-----------------------------|---------|---------|--|--|
| Fuerstinnenkorrespondenz    | 600     | 262.465 |  |  |
| RIDGES_Herbology_Versior 29 | 154.267 |         |  |  |

# Dokumentenbrowser

## Beispiel RIDGES



Annotationen genutzt:

- dipl (Text)
- lb (Zeilenumbrüche)

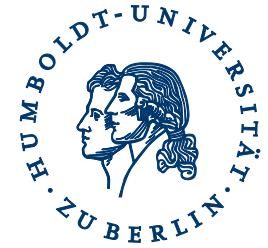
The screenshot shows the ANNIS web interface. At the top, there are navigation links: 'About ANNIS' and 'Report Problem'. A search bar contains 'Empty query'. Below the search bar, there are buttons for 'Search', 'More', and 'History'. A 'Corpus List' is visible, showing a table of corpora. The 'RIDGES\_Herbology\_Versior 29' corpus is selected and highlighted in blue. The main content area displays a preview of a document titled 'RIDGES\_Herbology\_Version4.1 > AlchemistischePraktik\_1603 - Visualizer: d'. The text is in German and discusses alchemical practices.

| Name                        | Texts | Tokens  |  |  |
|-----------------------------|-------|---------|--|--|
| Fuerstinnenkorrespondenz    | 600   | 262.465 |  |  |
| RIDGES_Herbology_Versior 29 |       | 154.267 |  |  |

RIDGES\_Herbology\_Version4.1 > AlchemistischePraktik\_1603 - Visualizer: d

An den Lefer .  
ES findt mir lieber Lefer / zwey feine Alchy-  
miftische Tractatlein von guten freundten  
zukomē / wiewol vnbeuwt wer fie zugerich-  
tet . Diefelben / weil ich fie zu allerley hand-  
griff zuverftehen / vnnd gute Artzneyen zube-  
reiten  
/ gut vnd nutz erachtet : hab ich dir fie nicht verhal-  
ten wöllen . Du magt davon vrtheilen / was du recht be-  
findeft . Ich hab mir das iudicium Athenienfium von  
Hippocrate belieben laffen / welchen fie vnder andern auch  
derenthalben zu ehren meineten : Quia fcripta de arte me-  
dica exacta edidit , multos effe volens medicos , qui ferua-  
re poſſent . Wuntſche auch daz folches Werck zu Gottes  
ehr / und deß Nechften fonderlich Deutſcher Nation , nutz  
wie es ahngeſehen / gereiche . Es ift nicht ohn / daz folche  
Deutſche werck den vngelehrten die kunft gemein machē /  
vnd alfo gleich den Sewen die Perlen vorwerffen . Iſt vor  
dieſē mein gemuth auch gewefen von folchen mich abzu-  
halten . Aber ich hab vermerckt / daz die gelehrten mehren  
theils der kunft nichts achten / vnd vielmehr darumb der  
publication zuwider fein / daz der gemeine Deutſche nicht  
verftehe wie gelehrt fie fein / vnnd alfo nicht dörffen fleißig  
der heimlichkeit der natur nachforſchen . Wolten lieber /

# Annotationsanzeige



3 Path: shenoute.a22 > a22.YA421-428 (norm\_group 52 - 75) left context: 5

ολ εγεχνεπρλλο αγω χωρισται ρενωρη φνη η σελααγ ηρωμε εγφωανσρατογ ερραι ρηλααγ ηρηλααγ ητειμινε εγεφωπε εγρογορτ

<

- ⊞ annotations (grid)
- ⊞ analytic view (document)
- ⊞ diplomatic text (document)

|                |               |
|----------------|---------------|
| ρμηεραϊεγ      | περιτογερη    |
| βαγκλιον       | τεηογα ενη    |
| ηεραϊεγω       | σομημοχ ε     |
| φωγεραϊ        | μοσφε ει      |
| 5 εσελααγηρηνα | 5 εβολ εγεχνε |
| αγ εφτητων     | πρλλοαγω      |
| εναϊεμπογ      | Χωρισταιρεν   |
| ογερσαρνε      | φρηωρηη       |
| ναχητηηρηλ     | ησελααγηρω    |
| 10 λωειμητι    | 10 μεεγφωανσρ |
| πηαηηρωμε      | ρατογεραηρη   |
| εφωμεη         | λααγηρηλααγ   |
| ματηηηρηλ      | ητειμινεεγ    |
| λο αγρηρηη     | εφωπε εγ      |
| 15 ροηπεαγω    | 15 ερογορτ    |
| νεικοογε ε     | φρογορτη[σι]  |
| τρεφωηη        | πεηηατωφ      |

Annotationsanzeige

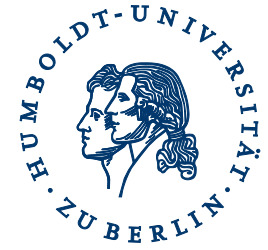
Treffer

Beispiel aus dem Korpus shenoute.a22:  
**urn:cts:copticLit:shenoute.a22**

Projekt Coptic Scriptorium

[https://corpling.uis.georgetown.edu/anis/scriptorium#\\_q=cDoicCI&c=c2hlm91dGUuYTly&cl=5&cr=5&s=0&l=10&\\_seg=bm9ybV9ncm91cA](https://corpling.uis.georgetown.edu/anis/scriptorium#_q=cDoicCI&c=c2hlm91dGUuYTly&cl=5&cr=5&s=0&l=10&_seg=bm9ybV9ncm91cA)

vgl. Zeldes und Schroeder (2015)



# MULTIPLE SEGMENTIERUNGEN

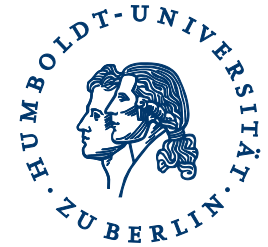
# Textebenen



- historische Korpora beinhalten u.a. häufig verschiedene Textebenen, z.B.
  - Transkriptionen
  - Normalisierungen
  - Übersetzungen
- Prinzip der multiplen Segmentierungen (Krause et al. 2012)
  - erlaubt unabhängige Segmentierungen
  - potentiell unendlich viele Textebenen
  - Dokumentation unter [http://corpus-tools.org/annis/resources/ANNIS3\\_multiseg\\_guide\\_2013-6.pdf](http://corpus-tools.org/annis/resources/ANNIS3_multiseg_guide_2013-6.pdf)

# Segmentierungen

## RIDGES



|              |     |     |      |        |                 |         |      |
|--------------|-----|-----|------|--------|-----------------|---------|------|
| <b>clean</b> | die | aus | drei | Worten | zusammengesetzt |         | sind |
| <b>dipl</b>  | die | aus | drei | Worten | zusammen-       | gesetzt | find |
| <b>lb</b>    | lb  |     |      |        |                 | lb      |      |
| <b>lemma</b> | d   | aus | drei | Wort   | zusammensetzen  |         | sein |
| <b>norm</b>  | die | aus | drei | Worten | zusammengesetzt |         | sind |

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

Trefferreferenz: <https://korpling.org/annis3/?id=f4d08b31-e1cd-434f-a81b-ebcoa325f21e>

# Segmentierungen

## RIDGES



| Position dipl | 1   | 2   | 3    | 4      | 5               | 6       | 7    |
|---------------|-----|-----|------|--------|-----------------|---------|------|
| Position norm | 1   | 2   | 3    | 4      | 5               |         | 6    |
| <b>clean</b>  | die | aus | drei | Worten | zusammengesetzt |         | sind |
| <b>dipl</b>   | die | aus | drei | Worten | zufammen-       | gefetzt | find |
| <b>lb</b>     | lb  |     |      |        |                 | lb      |      |
| <b>lemma</b>  | d   | aus | drei | Wort   | zusammensetzen  |         | sein |
| <b>norm</b>   | die | aus | drei | Worten | zusammengesetzt |         | sind |

Anzahl der Einheiten  
dipl=7  
norm=6

| Position dipl | 1   | 2             | 3    | 4     | 5          |          |
|---------------|-----|---------------|------|-------|------------|----------|
| Position norm | 1   | 2             | 3    | 4     | 5          | 6        |
| <b>clean</b>  | Den | Winter-spinet | sehr | groz  | zubekommen |          |
| <b>dipl</b>   | Den | Winter-fpinet | fehr | grofz | zubekommen |          |
| <b>lb</b>     | lb  |               |      |       |            |          |
| <b>lemma</b>  | d   | Winterspinat  | sehr | groß  | zu         | bekommen |
| <b>norm</b>   | Den | Winterspinat  | sehr | groß  | zu         | bekommen |

Anzahl der Einheiten  
dipl=5  
norm=6

Trefferreferenz: <https://korpling.org/annis3/?id=6576bc60-c437-48f4-b884-bfb3cc1do62d>

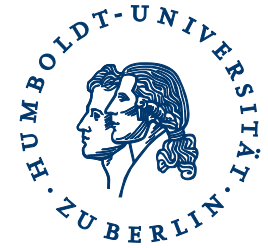
Trefferreferenz: <https://korpling.org/annis3/?id=f4do8b31-e1cd-434f-a81b-ebcoa325f21e>



# ANNOTATIONSANZEIGEN



# Annotationskonzepte



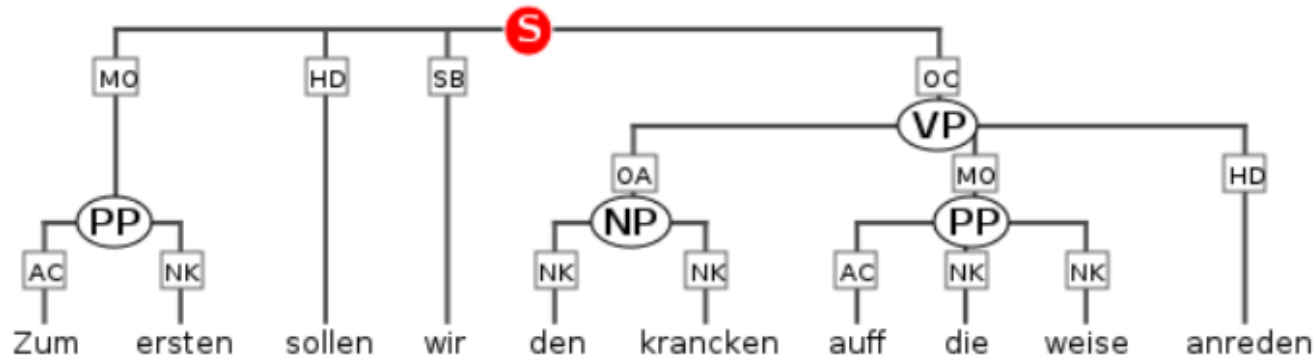
- verschiedene Annotationskonzepte erfordern häufig eine spezielle Anzeige
  - oft den Annotationstools und ihren Anzeigen nachempfunden
  - erleichtern die Analyse des Korpus in ANNIS
  - RIDGES und Fürstinnenkorrespondenzkorpus als Beispiel für Annotationsanzeigen von Token und Spannen
- hier nur ein paar Beispiele
  - unter <http://corpus-tools.org/annis/visualizations.html> alle Anzeigen für Annotationskonzepte (wie z.B. für Parallelkorpora, Multimediakorpora)

# Annotationskonzept Baum



- Beispiel aus: Hirschmann, Hagen; Linde, Sonja; DDB (Version 1.0), Humboldt-Universität zu Berlin. <http://korpling.german.hu-berlin.de/ddb-doku/index.htm>.  
<http://hdl.handle.net/11022/0000-0000-2107-3>
- vgl. Hirschmann und Linde (2010)

[\\* Show in ANNIS search interface](#)



Trefferreferenz: <https://korpling.org/annis3/?id=221a71ba-0a4b-4554-9969-b427d9cf136a>

# Annotationskonzept



- Beispiel aus dem Korpus: Ágel, Vilmos; Hennig, Mathilde; KAJUK (Version 1.1), Justus-Liebig-Universität Gießen. <http://www.uni-giessen.de/kajuk/index.htm>.  
<http://hdl.handle.net/11022/0000-0000-2102-8>
- vgl. Ágel und Hennig (2007)

[\\* Show in ANNIS search interface](#)

ein Herzog von Weimar . der Herzog von Weimar War **aber** gestorben und führt sie ein

**Co Component: 1, Type: E**

Trefferreferenz: <https://korpling.org/annis3/?id=56a854f1-67f5-4121-8085-15a8426ebf4c>



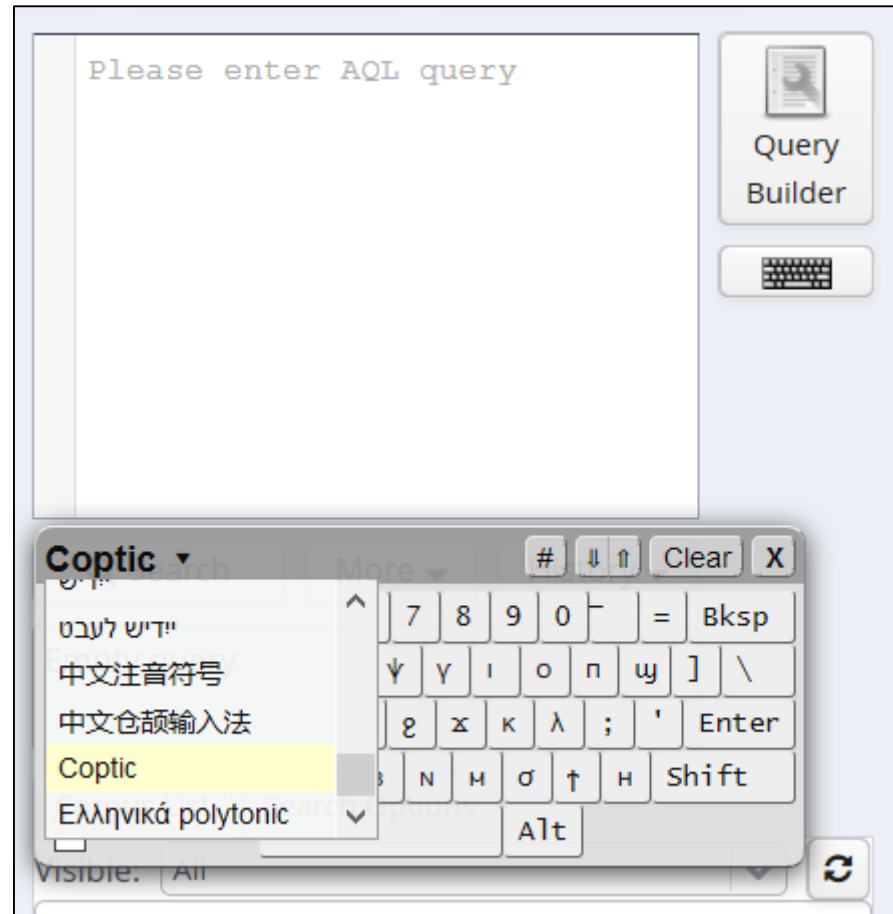
# UNTERSTÜTZUNG VON SCHRIFTARTEN

# Schriftarten

## Virtuelles Keyboard



- Einbinden verschiedener Schriftarten in ANNIS
  - z.B. Chinesisch, Koptisch, Kyrillisch oder Arabisch



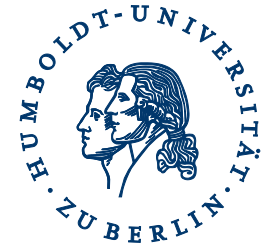


# ANNIS-Tutorials



- erste Tutorialvideos zum
  - Interface
  - Annotationsarten und -anzeigen
  - Suche nach Wortartenannotation („PoS“)
  - unter <https://www.linguistik.hu-berlin.de/de/institut/professuren/korpuslinguistik/corpus-tools/annis-tutorials>
- Weitere Videos sind geplant!
- User-Dokumentation: [http://corpus-tools.org/annis/resources/ANNIS User Guide 3.3.5.pdf](http://corpus-tools.org/annis/resources/ANNIS_User_Guide_3.3.5.pdf)
- Wünsche und Anmerkungen an [korpling@hu-berlin.de](mailto:korpling@hu-berlin.de)

# Kooperation



- unterstützt durch

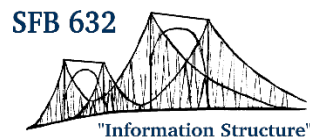
Humboldt-Universität zu Berlin

Georgetown University Washington



Sonderforschungsbereich 632

Deutsche Forschungsgemeinschaft



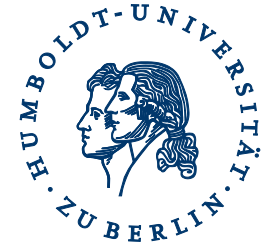
und viele weitere!

VIELEN DANK FÜR IHRE AUFMERKSAMKEIT!





# Software



## ANNIS Such- und Visualisierungstool

- Homepage: <http://corpus-tools.org/annis>
- HU-Instanz, Suche: <https://korpling.german.hu-berlin.de/annis3/>

## PEPPER Konverterframework

- Homepage: <http://corpus-tools.org/pepper>
- Technische Dokumentation: <https://github.com/korpling/pepper>

# Korpora



- Deutsche Diachrone Baumbank
  - Hirschmann, Hagen; Linde, Sonja; DDB (Version 1.0), Humboldt-Universität zu Berlin.  
<http://korpling.german.hu-berlin.de/ddb-doku/index.htm>.  
<http://hdl.handle.net/11022/0000-0000-2107-3>
- Fürstinnenkorrespondenzkorpora
  - Lühr, Rosemarie; Faßhauer, Vera; Prutscher, Daniela; Seidel, Henry; Fürstinnenkorrespondenz (Version 1.1), Universität Jena, DFG. <http://www.indogermanistik.uni-jena.de/Web/Projekte/Fuerstinnenkorr.htm>.  
<http://hdl.handle.net/11022/0000-0000-82A0-7>
- Kasseler Junktionskorpora
  - Ágel, Vilmos; Hennig, Mathilde; KAJUK (Version 1.1), Justus-Liebig-Universität Gießen.  
<http://www.uni-giessen.de/kajuk/index.htm>.  
<http://hdl.handle.net/11022/0000-0000-2102-8>
- Register in Diachronic German Science
  - Lüdeling, Anke; Odebrecht, Carolin; Zeldes, Amir; RIDGES-Herbology (Version 4.1), Humboldt-Universität zu Berlin. <http://korpling.german.hu-berlin.de/ridges/>.  
<http://hdl.handle.net/11022/0000-0000-8253-F>
- Shenoute.a22
  - urn:cts:copticLit:shenoute.a22, Projekt Coptic Scriptorium (Amir Zeldes, Caroline T. Schroeder)  
[https://corpling.uis.georgetown.edu/annis/scriptorium#\\_c=c2hlbm91dGUuYTly](https://corpling.uis.georgetown.edu/annis/scriptorium#_c=c2hlbm91dGUuYTly)

# Referenzen



- **Ágel, Vilmos; Hennig, Mathilde** (2007): DFG-Projekt “Explizite und elliptische Junktion in der Syntax des Neuhochdeutschen” Forschungsnotiz. In: Zeitschrift für Germanistische Linguistik 35, 185-189.
- **Belz, Malte; Odebrecht, Carolin; Perlitz, Laura, Voigt, Vivian** (2015): Annotationsrichtlinien zu Ridges Herbology Version 4.1, Humboldt-Universität zu Berlin. [http://korpling.german.hu-berlin.de/ridges/download/pubs/annotationGuidelines\\_v4.1.pdf](http://korpling.german.hu-berlin.de/ridges/download/pubs/annotationGuidelines_v4.1.pdf) [Letzter Zugriff am 05.01.2016].
- **Hirschmann, Hagen; Linde, Sonja** (2010): Annotationsguidelines zur Deutschen Diachronen Baumbank. Technical Report. Humboldt-Universität zu Berlin. [https://www.linguistik.hu-berlin.de/de/institut/professuren/korpuslinguistik/mitarbeiterinnen/hagen/ddb\\_doku](https://www.linguistik.hu-berlin.de/de/institut/professuren/korpuslinguistik/mitarbeiterinnen/hagen/ddb_doku) [letzter Zugriff am 01.02.2016].
- **Krause, Thomas; Zeldes, Amir** (2014): ANNIS3: A new architecture for generic corpus query and visualization. in: Digital Scholarship in the Humanities 2014 <http://dsh.oxfordjournals.org/cgi/content/abstract/fqu057?ijkey=GJBr0LhNfKW1g8i&keytype=ref> [letzter Zugriff: 20.10.15]
- **Krause, Thomas; Lüdeling, Anke; Odebrecht, Carolin; Zeldes, Amir** (2012): Multiple Tokenization in a Diachronic Corpus. In Exploring Ancient Languages through Corpora Conference, EALC 2012. Oslo.
- **Odebrecht, Carolin; Belz, Malte; Zeldes, Amir & Lüdeling, Anke** (eingereicht) RIDGES Herbology - Designing a Diachronic Multi-Layer Corpus. <https://www.linguistik.hu-berlin.de/institut/professuren/korpuslinguistik/mitarbeiterinnen/anke/pdf/odebrechtetalridges-submitted.pdf> [letzter Zugriff am 01.02.2016].
- **Zeldes, Amir ; Schroeder, Caroline T.** (2015): Computational Methods for Coptic: Developing and Using Part-of-Speech Tagging for Digital Scholarship in the Humanities. *Digital Scholarship in the Humanities* 31(1), 164-176. [https://corpling.uis.georgetown.edu/amir/pdf/Computational\\_Methods\\_for\\_Coptic\\_prepub.pdf](https://corpling.uis.georgetown.edu/amir/pdf/Computational_Methods_for_Coptic_prepub.pdf) [letzter Zugriff 20.10.2015]
- **Zipser, Florian; Romary, Laurent** (2010): A Model Oriented Approach to the Mapping of Annotation Formats using Standards. Proceedings of the Workshop on Language Resource and Language Technology Standards, LREC-2010 . Valletta, Malta, pp. 7–18.