Dokumentation und Annotationsrichtlinien für das Korpus Ridges Herbology Version 9.0 (ANNIS-und PAULA-Format*)

auf Grundlage des Metadatenframeworks nach LAUDATIO
Stand 31.03.2020
-deutsch-

Carolin Odebrecht
Gohar Schnelle
Catharina Fischer
Laura Perlitz

*Diese setzen sich aus Annotationen zusammen, die in verschiedenen Formaten erstellt wurden. Alle (Zwischen-) Formate stehen auf korpling.org/ridges zum Download bereit, enthalten aber jeweils nicht alle hier aufgeführten Annotationsebenen, bzw. einige zusätzliche (irrelevante) Ebenen, die nicht nach ANNIS und PAULA konvertiert wurden. Eine ausführlichere Dokumentation der einzelnen Formate finden Sie im LAUDATIO-Repositorium unter laudatio-repository.org.

Inhalt

1. Korpus 6
2. Dokumente 11
2.1. Kraeuterbuch_1914_Losch 11
2.2. NatürlichePflanzenfamilien_1887_Engler 11
2.3. GemeinnuetzigesKraeuterbuch_1874_Siegmund 12
2.4. FloraDerPreussischenRheinlande_1870_Wirtgen 12
2.5. DeutschePflanzennamen_1870_Grassmann 13
2.6. VorlesungenUeberKraeuterkunde_1843_Link 13
2.7. GrossesIllustriertesKraeuterbuch_1860_Mueller 14
2.8. NochEinigeWorte-VR_1840_Meyen 14
2.9. NochEinigeWorte_1840_Meyen 15
2.10. EigenschaftenAllerHeilpflanzen_1828_Anonymous 15
2.11. EigenschaftenAllerHeilpflanzen-149-165_1828_Anonymous 16
2.12. Flora-7_1821_Sieber 16
2.13. Flora-6_1821_Wilbrand 17
2.14. GetreueDarstellungUndBeschreibung_1809_Hayne 17
2.15. Apothekerlexikon_1793_Hahnemann 18
2.16. GrundrissKraeuterkunde_1792_Willdenow 18
2.17. Unterricht_1774_Eisen 19
2.18. BeschreibungderGraeser_c27-31_1769_Schrebers 19
2.19. BeschreibungderGraeser_c37-48_1769_Schrebers 20
2.20. EinleitungZuDerKraeuterkenntnisz_1764_Oeder 20
2.21. BlackwellischesKraeuterbuch_1750_Blackwell 21
2.22. MysterivmSigillorvm_1735_Hiebner 21
2.23. FloraSaturnizans_1722_Henckel 22
2.24. ViridariumReformatum_1719_Valentini 22
2.25. TheatrumBotanicum_1696_Verzascha 23
2.26. SchweizerischerBotanicus_1687_vonRoll 23
2.27. SonderbaresKraeuterbuch-21-36_1675_Anonymous 24
2.28. SonderbaresKraeuterbuch-11-21_1675_Anonymous 25
2.29. SonderbaresKraeuterbuch-1-11_1675_Anonymous 25
2.30. ThesaurusSanitatis_304-321_1673_Nasser 26
2.31. ThesaurusSanitatis_1673_Nasser 26
2.32. Phythologia_1662_Becher 27

2.33. Wund-Artzney_1652_Greiff ... 27
2.34. PflantzGart-VR_1639_Rhagor ... 28
2.35. PflantzGart-c4_1639_Rhagor ... 28
2.36. PflantzGart_1639_Rhagor ... 29
2.37. Kraeuterbuch_1609_Carrichter ... 29
2.38 Kraeutterbuch-VR_1609_Carrichter ... 30
2.39. HortulusSanitatis_1609_Uffenbach ... 30
2.40. HortulusSanitatis-VR_1609_Uffenbach ... 31
2.41. AlchymistischePractic_1603_Libavius ... 31
2.42. AlchymistischePractic-VR_1603_Libavius ... 32
2.43. NeuwKreuterbuch_1588_Theodorus ... 32
2.44. NeuwKreuterbuch-VR_1588_Theodorus ... 33
2.45. Paradeiszgaertlein_1588_Rosbach ... 33
2.46. Paradeiszgaertlein-VR_1588_Rosbach ... 34
2.47. NewKreueterbuch_1563_Handsch ... 34
2.48. NewKreuterbuch-VR_1563_Handsch ... 35
2.49. WieSichMeniglich_1557_vonBodenstein ... 35
2.50. WieSichMeniglich-VR_1557_vonBodenstein ... 36
2.51. Experimenta_1550_Schellenberg ... 36
2.52. Experimenta-VR_1550_Schellenbeg ... 37
2.53. NewKreueterbuch_1543_Fuchs ... 37
2.54. NewKreuterbuch-c19-25_1543_Fuchs ... 38
2.55. NewKreuterbuch-VR_1543_Fuchs ... 38
2.56. NewKreuetterBuch_1539_Bock ... 39
2.57. NewKreuetterbuch-CXXXIIIJ-CXXXVIIJ_1539_Bock ... 39
2.58. NewKreuetterBuch-VR_1539_Bock ... 40
2.59. ContrafaytKreuterbuch_1532_Brunfels ... 40
2.60. ContrafaytKreuterbuch- CCXXXVII-CCXLVIII_1532_Brunfels ... 41
2.61. ContrafaytKreuterbuch-CCLVII-CCLXX_1532_Brunfels ... 41
2.62. ContrafaytKreuterbuch-CCLXXI-CCXC_1532_Brunfels ... 42
2.63. ContrafaytKreuterbuch-VR_1532_Brunfels ... 43
2.64. ArtzneyBuchleinDerKreutter_1532_Tallat ... 43
2.65. ArtzneyBuchleinDerKreutter-Abrotanum_1532_Tallat ... 44
2.66. ArtzneyBuchleinDerKreutter-Cretanus_1532_Tallat ... 44
2.67. ArtzneyBuchleinDerKreutter-VR_1532_Tallat ... 45

2.68. GartDerGesundheit_1487_vonCuba ... 45
2.69. GartDerGesundheit-c15-23_1487_Cuba ... 46
2.70. GartDerGesundheit-c23-33_1487_Cuba ... 46
2.71. GartDerGesundheit-VR_1487_vonCuba ... 47
2.72. BuchDerNatur_1482_vonMegenberg ... 47
2.73. BuchDerNatur-Canamellis_1482_vonMegenberg ... 48
3. Annotationsebenen – Transkription/Normalisierung ... 48
3.1. dipl ... 48
3.2. clean ... 50
3.3. norm ... 52
4. Annotationsebenen – Annotationen zu linguistischen Eigenschaften ... 55
4.1. pos ... 55
4.2. lemma ... 57
4.3. lang ... 59
4.4. abbr ... 61
4.5. comp ... 63
4.6. comp_orth ... 64
4.7. prot ... 67
4.8. comp_n ... 69
4.9. comp_n_graph ... 71
4.10. comp_n_mod ... 73
4.11. comp_a ... 75
4.12. comp_a_graph ... 77
4.13. comp_amb ... 79
4.14. infl_fuge ... 81
4.15. comp_lex ... 83
4.16. adja_uninfl ... 85
4.17. attr_gen ... 87
4.18. morph_ellipsis ... 89
4.19. kaus1 ... 90
4.20. kaus2 ... 92
4.21. kaus3 ... 93
4.22. kaus4 ... 95
4.23. kon1 ... 96
4.24. kon2 ... 98

4.25. lin1 99
4.26. lin2 101
4.27. lin3 102
4.28. lin4 104
4.29. hypo 105
4.30. M 107
4.31. gram 108
4.32. technicalterm 110
5. Annotationsebenen – Strukturelle Annotation 111
5.1. lb 111
5.2. quotemark 113
5.3. column 115
5.4. pb 116
5.5. pb_n 118
5.6. unclear 120
5.7. figure 122
5.8. figure_p 132
5.9. script 134
5.10. hi_rend 136
5.11. head 138
5.12. note 140
5.13. quote 142
6. Annotationsebenen – Inhaltliche Annotation 143
6.1. persname 144
6.2. title 148
6.3. author_ref 149
6.4. reader_ref 151
6.5. Metadaten 153
7. Referenzen 156
Anhang 157
Zuordnung von Annotationen und Dokumenten 157
Transkription und Normalisierung 170
In Verbindung mit Excel benutzte Tools 193

1. Korpus
Typ: Corpus – Ridges Herbology

Name:	Ridges Herbology
Herausgeber:	Anke Lüdeling, Carolin Odebrecht, Thomas Krause, Gohar Schnelle, Catharina Fischer Korpuslinguistik, Humboldt-Universität zu Berlin
Annotatoren:	Ilham Abed-Ali, Diana Allamirano, Silke Andresen, Henriette Ast, Valeska Beckert, Malte Belz, Katrin Benning, Thomas Berg, Cornelia Binnyus, Simon Blum, Swenja Bödecker, Doreen Christen, Camilla Colombo, Janin Czilwik, Bianca David, Mascha Dayal, Antonia Dittberner, Cora Döhn, Imke Driemel, Elisabeth Eberle, Christian Ebert, Olja Efremova, Gill-Maria Eichhorn, Stefanie Enneper, Judith Esser, Beate Federau, Nico Friesenhan Catharina Fischer, Karolina Flisikowski, Nerina Gärtner, Felix Gehrke, Selina Gellweiler, Annegret Gerlach, Julia Gerolf, Linda Giesel, Svenja Guhr, Rahel Gajaneh Hartz, Kyra Hellmann, Yelyzaveta Hiebert, Hagen Hirschmann, Morice Hömmen, Stephanie Jandt, Bodil Jessen, Maria Kampermann, Keqin Ji, Nikolaos Kartalis, Sebastian Kiraga, Laura Köhler, Ewa Anna Kolbik, Birka Kondeyne, Anna Konik, Kornél Kovàcs, Marco Krämer, Daisy Krüger, Svetlana Kucenko, Till Kulawik, Anna-Maria Lehmann, Eliese-Sophia Lincke, Maria Lober, Robert Lorenz, Anke Lüdeling, Laura Lueders, Tatjana Malon, Samuele Maniscalco, Maria Martynova, Kim Kristin Maser, Laura McKee, Manuel Metzig, Alexander Meyer, Annika Mittelstädt, Lisa Mohs, Sandra Müller, Vinzent Müller, Andrew Murphy, Johannes Mursell, Juliane Nau, Carolin Odebrecht, Akiko Okuda, Jonas Opitz, Mareike Otto, Maik Papenhagen, Angelo Papenhoff, Laura Perlitz, Tilo Petersohn, Simona Petrova, Viktoria Peselmann, Steffen Pfennigwerth, Joachim Polte, Gina Reetz, Valerie Reichardt, Katharina Reinig, Guendalina Reul, Ina Riesler, Lena Rosin, Romy Sachs, Franz-Josef Sachse, Anna Sapronova, Simon Sauer, Jan Christian Schaffert, Anna Greta Schmahl, Claudia Schmidt, Gohar Schnelle, Johannes Schütz, Athina Sioupi, Sarah Sippach, Andrea Skotovic, Lucia Sohmen, Iryna Sorokovska, Madeleine Spitzer, Uwe Springmann, Kristina Stephan, Helen Stuke, Aleksandra Swiech, Franka Szperlinski, Juliane Tiemann, Anna Tóth, Réka Tóth, Phuong Thao Tran, Alexander Turtureanu, Gertje Utpatel, Esra Uyanik, Hanna Varachkina, Magdalena Vock, Mirjam Vogler, Marike de Vries Monika Walak, Linnéa Clara Weitkamp, Andrea Weißbarth, Juliana Wekel, Vita Rosalie Wijffels, Alexandra Wiemann, Svenja Wolter, Isabell Ines Zander, Karolina Zuchewicz (nachfolgend Studierende der Humboldt-Universität zu Berlin und der Georg-August-Universität Göttingen genannt)
Veröffentlichung:	31.03.2019

Anzahl der Textausschnitte:	73
Zeitraum:	1482-1914
dipl-Token:	305056
Sprache:	Deutsch
Register:	Kräutertexte
Zugang:	Suche und Visualisierung mit ANNIS3 (siehe Krause et al. (2016)): https://korpling.german.hu-berlin.de/annis3/
Projekt:	http://korpling.org/ridges
Projektbeschreibung:	Das **RIDGES**-Projekt (**R**egister **i**n **D**iachronic **Ge**rman **S**cience) untersucht die Entstehung und Entwicklung der deutschen Wissenschaftssprache ab Mitte des 15. Jahrhunderts bis ins späte 19. Jahrhundert. Bis zum 16. Jahrhundert war die Sprache der Wissenschaft in Europa vorwiegend Latein und erst im Laufe des 15. Jahrhunderts beginnen deutsche Wissenschaftler, wissenschaftliche Texte (oder zumindest Texte mit wissenschaftlichen Inhalten für Laien) erstmalig auch auf Deutsch zu formulieren. Dazu mussten sie ein wissenschaftliches Register ‚erfinden' – sie mussten die Terminologie genauso wie angemessene Textstrukturen entwickeln und erproben. Wissenschaftliche Texte haben sich über die Jahrhunderte natürlich verändert. Dabei sind die meisten Veränderungen quantitativ und nicht qualitativ (kategorial). Man kann die Entwicklungen nur nachzeichnen, wenn man Texte oder Textsammlungen von vielen Zeitpunkten genau miteinander vergleicht. In unserem Projekt sollen wissenschaftliche Texte daher auf allen sprachlichen Ebenen (Syntax, Wortbildung, Lexik, Phraseologie, Textstruktur etc.) analysiert werden, um Entwicklungen und Tendenzen identifizieren und beschreiben zu können. In einem variationistischen Ansatz annotieren wir wissenschaftliche Texte mit korpuslinguistischen Methoden und nutzen quantitative Verfahren, um Veränderungen in den Merkmalen zu identifizieren. Wichtig ist uns dabei die Mitwirkung unserer Studierenden, die im Rahmen ihres Studiums bei der Erstellung und der Annotation der Korpus-Ressourcen mit einbezogen werden. Die so entstandenen Ressourcen werden in mehreren Formaten unter einer Creative-Commons-Lizenz frei zur Verfügung gestellt. Das RIDGES-Projekt wird von einem Google Digital Humanities Research Award unterstützt. Archiviert ist das Koprus im LAUDATIO-Repository: www.laudatio-repository.org
Forschungsfrage:	In unserem Projekt sollen wissenschaftliche Texte auf allen sprachlichen Ebenen (Syntax, Wortbildung, Lexik, Phraseologie, Textstruktur etc.) analysiert werden, um Entwicklungen und

	Tendenzen identifizieren und beschreiben zu können. In einem variationistischen Ansatz annotieren wir wissenschaftliche Texte mit korpuslinguistischen Methoden und nutzen quantitative Verfahren, um Veränderungen in den Merkmalen zu identifizieren.
Version:	9.0
Korpusarchitektur:	Das diachrone Korpus Ridges Herbology besitzt eine multiple Segmentierung, die es erlaubt, konfligierende Tokenisierung von multiplen Ebenen zu verarbeiten (siehe Odebrecht et al. (2017)).
Art der Änderung:	Folgende Dokumente wurden in Version 9.0 hinzugefügt: BuchDerNatur-Canamellis_1482_vonMegenberg NewKreuterbuch-VR_1543_Fuchs Experimenta_1550_Schellenberg Experimenta-VR_1550_Schellenbeg NewKreuterbuch-VR_1563_Handsch NeuwKreuterbuch_1588_Theodorus Paradeiszgaertlein-VR_1588_Rosbach NeuwKreuterbuch-VR_1588_Theodorus HortulusSanitatis-VR_1609_Uffenbach Kraeutterbuch-VR_1609_Carrichter BeschreibungDerGraeser_c27-31_1769_Schrebers BeschreibungDerGraeser_c37-48_1769_Schrebers außerdem: - Durchgängige Korrektur der lang-Annotationen (nach erweiterten Annotationsrichtlinien) -In den Dokumenten HortulusSanitatis-VR_1609_Uffenbach und BeschreibungDerGraeser-c27-31_1769_Schrebers wurde die Annotationsebene „quotemark“ hinzugefügt -In allen Dokumenten wurden die Annotationsebenen „abbr“ , „gram“, „technicalterm“ und „M“ hinzugefügt -In folgenden Dokumenten wurden die Annotationsebenen „kaus1“, „kaus2“, „kaus3“, „kaus4“, „kon1“, „kon2“, „lin1“, „lin2“, „lin3“, „lin4“ und „hypo“ hinzugefügt: ArtzneyBuchleinDerKreutter-Abrotanum_1532_Tallat ArtzneyBuchleinDerKreutter-VR_1532_Tallat

GartDerGesundheit-c23-33_1487_Cuba
GartDerGesundheit-VR_1487_vonCuba
PflantzGart_1639_Rhagor
PflantzGart-VR_1639_Rhagor

-In dem Dokument NeuwKreuterbuch_1588_Theodorus konnten die Bilder der Seiten 34-36 nicht annotiert werden

-Folgende Annotationsebenen wurden (soweit vorhanden) aus allen Dokumneten erntfernt: orc, cat, cat-const, func, deprel, lemma-deprel, pos-deprel, morph, pos-const, foreign, foreign_trans, sentence_end, position_verb, subclause_type, KOUS_sem, comment_lex, comment, reference, form_disease, problem, herbname_norm, herbprep, form_prep, noun_nom, form_noun, clause_type, position_rel, position_referent, form_referent, position_verb_rel, form_relpron, mod_referent, ppk_e1, ppk_e2, ppk_e3, pb_ana, atLeast, atMost, interpretation, hi, ref, ref_target, ref_type, definition, disease, plant, property, name, name_type

Enthaltende Dokumente (Kürzel):

BuchDerNatur_1482_vonMegenberg
GartDerGesundheit_1487_vonCuba
GartDerGesundheit-c15-23_1487_Cuba
GartDerGesundheit-c23-33_1487_Cuba
GartDerGesundheit-VR_1487_vonCuba
ArtzneyBuchleinDerKreutter_1532_Tallat
ArtzneyBuchleinDerKreutter-Abrotanum_1532_Tallat
ArtzneyBuchleinDerKreutter-Centaurea_1532_Tallat
ArtzneyBuchleinDerKreutter-VR_1532_Tallat
ContrafaytKreuterbuch_1532_Brunfels
ContrafaytKreuterbuch-CCLVII-CCLXX_1532_Brunfels
ContrafaytKreuterbuch-CCLXXI-CCXC_1532_Brunfels
ContrafaytKreuterbuch-VR_1532_Brunfels
ContrafaytKreuterbuch-CCXXXVII-CCXLVIII_1532_Brunfels
NewKreuetterBuch_1539_Bock
NewKreuetterbuch-CXXXIIIJ-CXXXVIIJ_1539_Bock
NewKreuetterBuch-VR_1539_Bock
NewKreuterbuch_1543_Fuchs
NewKreuterbuch-c19-25_1543_Fuchs
NewKreuterbuch-VR_1543_Fuchs
WieSichMeniglich_1557_vonBodenstein
WieSichMeniglich-VR_1557_vonBodenstein
NewKreueterbuch_1563_Handsch
NewKreuterbuch-VR_1563_Handsch
Experimenta_1550_Schellenberg
Experimenta-VR_1550_Schellenbeg
NeuwKreuterbuch_1588_Theodorus

NeuwKreuterbuch-VR_1588_Theodorus
Paradeiszgaertlein_1588_Rosbach
Paradeiszgaertlein-VR_1588_Rosbach
AlchymistischePractic_1603_Libavius
AlchymistischePractic-VR_1603_Libavius
HortulusSanitatis_1609_Uffenbach
HortulusSanitatis-VR_1609_Uffenbach
Kraeutterbuch_1609_Carrichter
Kraeutterbuch-VR_1609_Carrichter
PflantzGart-c4_1639_Rhagor
PflantzGart_1639_Rhagor
PflantzGart-VR_1639_Rhagor
Wund-Artzney_1652_Greiff
Phythologia_1662_Becher
ThesaurusSanitatis_1673_Nasser
ThesaurusSanitatis_304-321_1673_Nasser
SonderbaresKraeuterbuch-1-11_1675_Anonymous
SonderbaresKraeuterbuch-11-21_1675_Anonymous
SonderbaresKraeuterbuch-21-36_1675_Anonymous
SchweizerischerBotanicus_1687_vonRoll
TheatrumBotanicum_1696_Verzascha
ViridariumReformatum_1719_Valentini
FloraSaturnizans_1722_Henckel
MysterivmSigillorvm_1735_Hiebner
BlackwellischesKraeuterbuch_1750_Blackwell
EinleitungZuDerKraeuterkenntnisz_1764_Oeder
BeschreibungDerGraeser_c27-31_1769_Schrebers
BeschreibungDerGraeser_c37-48_1769_Schrebers
Unterricht_1774_Eisen
GrundrissKraeuterkunde_1792_Willdenow
Apothekerlexikon_1793_Hahnemann
GetreueDarstellungUndBeschreibung_1809_Hayne
Flora-6_1821_Wilbrand
Flora-7_1821_Sieber
EigenschaftenAllerHeilpflanzen_1828_Anonymous
EigenschaftenAllerHeilpflanzen-149-165_1828_Anonymous
NochEinigeWorte_1840_Meyen
NochEinigeWorte-VR_1840_Meyen
VorlesungenUeberKraeuterkunde_1843_Link
GrossesIllustriertesKraeuterbuch_1860_Mueller
DeutschePflanzennamen_1870_Grassmann
FloraDerPreussischenRheinlande_1870_Wirtgen
GemeinnuetzigesKraeuterbuch_1874_Siegmund
NatürlichePflanzenfamilien_1887_Engler
Kraeuterbuch_1914_Losch

2. Dokumente

Alle Kurztitel werden, den Standards der ‚clean'-Normalisierung gemäß, nach Originalschreibweise realisiert, wobei Sonderzeichen, die der modernen Orthografie nicht entsprechen an den modernen Zeichensatz angepasst werden. Umlaute werden durch „ae", „ue" und „oe" ersetzt.

2.1. Kraeuterbuch_1914_Losch

Titel:	Kräuterbuch: Unsere Heilpflanzen in Wort und Bild
Kürzel:	Kraeuterbuch_1914_Losch
Autor:	Friedrich Losch
Jahr:	1914
Ort:	Eßlingen, München
Verlag:	Schreiber
Seitenangabe:	163-168
Publikationstyp:	Monographie
dipl-Einheiten:	3961
Beschreibung:	Kräutertext
Annotationsebenen:	dipl, clean, norm, abbr, author_ref, figure, hi_rend, lang, lb, lemma, note, pb, pb_n, persname, pos, script, title

2.2. NatürlichePflanzenfamilien_1887_Engler

Titel:	Die natürlichen Pflanzenfamilien
Kürzel:	NatürlichePflanzenfamilien_1887_Engler
Autor:	Adolf Engler
Jahr:	1887
Ort:	Leipzig
Verlag:	Wilhelm Engelmann
Seitenangabe:	280-286
Publikationstyp:	Monographie
dipl-Einheiten:	3828
Beschreibung:	Kräutertext
Annotationsebenen:	dipl, clean, norm, abbr, author_ref, figure, head, hi_rend, lang, lb, lemma, pb, pb_n, persname, pos, quote, script

2.3. GemeinnuetzigesKraeuterbuch_1874_Siegmund

Titel:	Gemeinnütziges Kräuterbuch
Kürzel:	GemeinnuetzigesKraeuterbuch_1874_Siegmund
Autor:	Ferdinand Siegmund
Jahr:	1874
Ort:	Wien und Pest
Verlag:	Hartleben
Seitenangabe:	316-325
Publikationstyp:	Monographie
dipl-Einheiten:	4527
Beschreibung:	Kräutertext
Annotationsebenen:	dipl, clean, norm, abbr, author_ref, figure, hi_rend, lang, lb, lemma, pb, pb_n, persname, pos, quote, reader_ref, script

2.4. FloraDerPreussischenRheinlande_1870_Wirtgen

Titel:	Flora Der Preussischen Rheinlande
Kürzel:	FloraDerPreussischenRheinlande_1870_Wirtgen
Autor:	Wirtgen, Philipp Wilhelm
Jahr:	1870
Ort:	Bonn
Verlag:	NA
Seitenangabe:	1-22
Publikationstyp:	Monographie
dipl-Einheiten:	7519
Beschreibung:	Kräutertext
Annotationsebenen:	dipl, clean, norm, abbr, attr_gen, author_ref, comp, comp_orth, figure, head, hi_rend, lang, lb, lemma, morph_ellipsis, note, pb, pb_n, pos, quote, script

2.5. DeutschePflanzennamen_1870_Grassmann

Titel:	Deutsche Pflanzennamen
Kürzel:	DeutschePflanzennamen_1870_Grassmann
Autor:	Grassmann, Hermann
Jahr:	1870
Ort:	Stettin
Verlag:	NA
Seitenangabe:	1-23
Publikationstyp:	Monographie
dipl-Einheiten:	10282
Beschreibung:	Kräutertext
Annotationsebenen:	dipl, clean, norm, abbr, attr_gen, author_ref, comp, comp_orth, head, hi_rend, lang, lb, lemma, morph_ellipsis, pb, pb_n, pos, prot, quote, script, unclear

2.6. VorlesungenUeberKraeuterkunde_1843_Link

Titel:	Vorlesungen Über Kräuterkunde
Kürzel:	VorlesungenUeberKraeuterkunde_1843_Link
Autor:	Link , Heinrich Friedrich
Jahr:	1843
Ort:	Berlin
Verlag:	Verlag von C. G. Lüderitz
Seitenangabe:	28-48
dipl-Einheiten:	7995
Beschreibung:	Kräutertext
Annotationsebenen:	dipl, clean, norm, abbr, attr_gen, author_ref, comp, comp_orth, head, hi_rend, lang, lb, lemma, morph_ellipsis, note, pb, pb_n, pos, prot, quote, reader_ref, script

2.7. GrossesIllustriertesKraeuterbuch_1860_Mueller

Titel:	Das große illustrierte Kräuter-Buch
Kürzel:	GrossesIllustriertesKraeuterbuch_1860_Mueller
Autor:	Ferdinand Müller
Jahr:	1860
Ort:	Ulm
Verlag:	NA
Seitenangabe:	617-622
Publikationstyp:	Monographie
dipl-Einheiten:	3852
Beschreibung:	Kräutertext
Annotationsebenen:	dipl, clean, norm, abbr, author_ref, figure, head, hi_rend, lang, lb, lemma, pb, pb_n, persname, pos, quote, reader_ref, script

2.8. NochEinigeWorte-VR_1840_Meyen

Titel:	Noch einige Worte über den Befruchtungsakt und die Poljembryonie bei den höheren Pflanzen (Vorwort)
Kürzel	NochEinigeWorte-VR_1840_Meyen
Autor:	Meyen, Franz Julius Ferdinand
Jahr:	1840
Ort:	Berlin
Verlag:	Haude und Spenersche Buchhandlung (S. J. Joseephy.)
Seitenangabe:	III-IV
Publikationstyp:	Monographie
dipl-Einheiten:	365
Beschreibung:	Kräutertext
Annotationsebenen:	dipl, clean, norm, abbr, attr_gen, author_ref, comp, comp_orth, head, hi_rend, lang, lb, lemma, note, pb, pb_n, pos, prot, quote, script

2.9. NochEinigeWorte_1840_Meyen

Titel:	Noch einige Worte über den Befruchtungsakt und die Poljembryonie bei den höheren Pflanzen
Kürzel	NochEinigeWorte_1840_Meyen
Autor:	Meyen, Franz Julius Ferdinand
Jahr:	1840
Ort:	Berlin
Verlag:	Haude und Spenersche Buchhandlung (S. J. Joseephy.)
Seitenangabe:	24-50
Publikationstyp:	Monographie
dipl-Einheiten:	7619
Beschreibung:	Kräutertext
Annotationsebenen:	dipl, clean, norm, abbr, attr_gen, author_ref, comp, comp_orth, hi_rend, lang, lb, lemma, morph_ellipsis, note, pb, pb_n, pos, prot, quote, reader_ref, script

2.10. EigenschaftenAllerHeilpflanzen_1828_Anonymous

Titel:	Die Eigenschaften Aller Heilpflanzen
Kürzel	EigenschaftenAllerHeilpflanzen_1828_Anonymous
Autor:	NA
Jahr:	1828
Ort:	München
Verlag:	Jos. Lindauer'sche Buchhandlung
Seitenangabe:	3-21
dipl-Einheiten:	6849
Beschreibung:	Kräutertext
Annotationsebenen:	dipl, clean, norm, abbr, attr_gen, comp, comp_orth, hi_rend, lang, lb, lemma, morph_ellipsis, pb, pb_n, pos, prot, quote, script, unclear

2.11. EigenschaftenAllerHeilpflanzen-149-165_1828_Anonymous

Titel:	Die Eigenschaften Aller Heilpflanzen
Kürzel:	EigenschaftenAllerHeilpflanzen-149-165_1828_Anonymous
Autor:	NA
Jahr:	1828
Ort:	München
Verlag:	Jos. Lindauer'sche Buchhandlung
Seitenangabe:	149-165
Publikationstyp:	Monographie
dipl-Einheiten:	5561
Beschreibung:	Kräutertext
Annotationsebenen:	dipl, clean, norm, abbr, head, hi_rend, lang, lb, lemma, pb, pb_n, pos, quote, script

2.12. Flora-7_1821_Sieber

Titel:	Flora oder Botanische Zeitung 7
Kürzel:	Flora-7_1821_Sieber
Autor:	Sieber
Jahr:	1821
Ort:	Regensburg
Verlag:	Königl. botanische Gesellschaft
Seitenangabe:	97-112
Publikationstyp:	Zeitschrift
dipl-Einheiten:	3636
Beschreibung:	Kräutertext
Annotationsebenen:	dipl, clean, norm, abbr, attr_gen, author_ref, comp, comp_orth, head, lang, lb, lemma, morph_ellipsis, pb, pb_n, pos, prot, quote, script, unclear

2.13. Flora-6_1821_Wilbrand

Titel:	Flora oder Botanische Zeitung 6
Kürzel:	Flora-6_1821_Wilbrand
Autor:	B. Wilbrand
Jahr:	1821
Ort:	Regensburg
Verlag:	Königl. botanische Gesellschaft
Seitenangabe:	81-96
Publikationstyp:	Zeitschrift
dipl-Einheiten:	3312
Beschreibung:	Kräutertext
Annotationsebenen:	dipl, clean, norm, abbr, attr_gen, author_ref, comp, comp_orth, head, lang, lb, lemma, note, pb, pb_n, pos, prot, quote, reader_ref, script, unclear

2.14. GetreueDarstellungUndBeschreibung_1809_Hayne

Titel:	Getreue Darstellung und Beschreibung der in der Arzneykunde gebräuchlichen Gewächse, wie auch solcher, welche mit ihnen verwechselt werden können
Kürzel:	GetreueDarstellungUndBeschreibung_1809_Hayne
Autor:	Friedrich Gottlob Hayne
Jahr:	1809
Ort:	Berlin
Verlag:	NA
Seitenangabe:	NA
Publikationstyp:	Monographie
dipl-Einheiten:	3177
Beschreibung:	Kräutertext
Annotationsebenen:	dipl, clean, norm, abbr, author_ref, figure, head, hi_rend, lang, lb, lemma, pb, persname, pos, quote, script, title, unclear

2.15. Apothekerlexikon_1793_Hahnemann

Titel:	Apothekerlexikon
Kürzel:	Apothekerlexikon_1793_Hahnemann
Autor:	Samuel Hahnemann
Jahr:	1793
Ort:	Leipzig
Verlag:	NA
Seitenangabe:	437-446
Publikationstyp:	Monographie
dipl-Einheiten:	4244
Beschreibung:	Kräutertext
Annotationsebenen:	dipl, clean, norm, abbr, head, hi_rend, lang, lb, lemma, pb, pb_n, pos, quote, script

2.16. GrundrissKraeuterkunde_1792_Willdenow

Titel:	Grundriss der Kräuterkunde zu Vorlesungen entworfen
Kürzel:	GrundrissKraeuterkunde_1792_Willdenow
Autor:	Willdenow, Carl Ludwig
Jahr:	1792
Ort:	Berlin
Verlag:	Haude und Spener
Seitenangabe:	282-305
Publikationstyp:	Monographie
dipl-Einheiten:	5094
Beschreibung:	Kräutertext
Annotationsebenen:	dipl, clean, norm, abbr, attr_gen, author_ref, comp, comp_orth, head, hi_rend, lang, lb, lemma, morph_ellipsis, pb, pb_n, pos, prot, quote, reader_ref, script, unclear

2.17. Unterricht_1774_Eisen

Titel:	Unterricht von der allgemeinen Kräuter- und Wurzeltrocknung
Kürzel:	Unterricht_1774_Eisen
Autor:	Eisen, Johann Georg
Jahr:	1774
Ort:	Riga
Verlag:	NA
Seitenangabe:	44-72
Publikationstyp:	Monographie
dipl-Einheiten:	3982
Beschreibung:	Kräutertext
Annotationsebenen:	dipl, clean, norm, abbr, attr_gen, author_ref, comp, comp_orth, head, hi_rend, lang, lb, lemma, morph_ellipsis, pb, pb_n, pos, prot, quote, reader_ref, script, unclear

2.18. BeschreibungderGraeser_c27-31_1769_Schrebers

Titel:	Beschreibung der Gräser nebst ihren Abbildungen nach der Natur
Kürzel:	BeschreibungderGraeser _c27-31_1769_Schrebers
Autor:	Schreber , Johann Christian Daniel
Jahr:	1769
Ort:	Leipzig
Verlag:	Siegfried Lebrecht Crusius
Seitenangabe:	27-31
Puplikationstyp:	Monographie
dipl-Einheiten:	2273
Beschreibung:	Kräutertext
Annotationsebenen:	dipl, clean, norm, abbr, author_ref, column, figure, figure_p, head, hi_rend, lang, lb, lemma, note, pb, pb_n, persname, pos, quotemark, reader_ref, script, title, unclear

2.19. BeschreibungderGraeser_c37-48_1769_Schrebers

Titel:	Beschreibung der Gräser nebst ihren Abbildungen nach der Natur
Kürzel:	BeschreibungderGraeser _c37-48_1769_Schrebers
Autor:	Schreber , Johann Christian Daniel
Jahr:	1769
Ort:	Leipzig
Verlag:	Siegfried Lebrecht Crusius
Seitenangabe:	37-48
Puplikationstyp:	Monographie
dipl-Einheiten:	5262
Beschreibung:	Kräutertext
Annotationsebenen:	dipl, clean, norm, abbr, author_ref, column, figure, figure_p, head, hi_rend, lang, lb, lemma, note, pb, pb_n, , persname, pos, reader_ref, script, title, unclear

2.20. EinleitungZuDerKraeuterkenntnisz_1764_Oeder

Titel:	Einleitung Zu Der Kräuterkenntnisz
Kürzel:	EinleitungZuDerKraeuterkenntnisz_1764_Oeder
Autor:	Oeder , George Christian
Jahr:	1764
Ort:	Kopenhagen
Verlag:	NA
Seitenangabe:	304-339
dipl-Einheiten:	7115
Beschreibung:	Kräutertext
Annotationsebenen:	dipl, clean, norm, abbr, attr_gen, author_ref, comp, comp_orth, head, hi_rend, lang, lb, lemma, morph_ellipsis, pb, pb_n, pos, prot, quote, reader_ref, script, unclear

2.21. BlackwellischesKraeuterbuch_1750_Blackwell

Titel:	Vermehrtes und verbessertes Blackwellisches Kräuter-Buch, das ist Elisabeth Blackwell Sammlung der Gewächse die zum Arzney-Gebrauch in den Apothecken aufbehalten werden, deren Beschreibung und Kräfften
Kürzel:	BlackwellischesKraeuterbuch_1750_Blackwell
Autor:	Elisaneth Blackwell
Jahr:	ca. 1750
Ort:	Nürnberg
Verlag:	Nicolaus Friedrich Eisenberger
Seitenangabe:	NA
Publikationstyp:	Monographie
dipl-Einheiten:	5609
Beschreibung:	Kräutertext
Annotationsebenen:	dipl, clean, norm, abbr, author_ref, head, hi_rend, lang, lb, lemma, pb, persname, pos, quote, script, title, unclear

2.22. MysterivmSigillorvm_1735_Hiebner

Titel:	Mysterivm Sigillorvm
Kürzel:	MysterivmSigillorvm_1735_Hiebner
Autor:	Hiebner von Schneebergk, Israel
Jahr:	1735
Ort:	Erfurt
Seitenangabe:	65-89
dipl-Einheiten:	7864
Beschreibung:	Kräutertext
Annotationsebenen:	dipl, clean, norm, abbr, attr_gen, author_ref, comp, comp_orth, head, hi_rend, lang, lb, lemma, morph_ellipsis, note, pb, pb_n, pos, position_verb, prot, quote, reader_ref, script, unclear

2.23. FloraSaturnizans_1722_Henckel

Titel:	Flora saturnizans, Die Verwandschafft des Pflanzen mit dem Mineral Reich
Kürzel:	FloraSaturnizans_1722_Henckel
Autor:	Henckel, Johann Friedrich
Jahr:	1722
Ort:	Leipzig
Verlag:	Johann Christian Martini
Seitenangabe:	647-671
Publikationstyp:	Monographie
dipl-Einheiten:	6219
Beschreibung:	Kräutertext
Annotationsebenen:	dipl, clean, norm, abbr, attr_gen, author_ref, comp, comp_orth, head, hi_rend, lang, lb, lemma, morph_ellipsis, note, pb, pb_n, pos, prot, quote, reader_ref, script, unclear

2.24. ViridariumReformatum_1719_Valentini

Titel:	Viridarium Reformatum, Seu Regnum Vegetabile, Das ist: Neu-eingerichtetes und Vollständiges Kräuter-Buch
Kürzel:	ViridariumReformatum_1719_Valentini
Autor:	Michael Bernhard Valentini
Jahr:	1719
Ort:	Frankfurt am Main
Verlag:	Heinscheidt
Seitenangabe:	459-465
Publikationstyp:	Monographie
dipl-Einheiten:	3596
Beschreibung:	Kräutertext
Annotationsebenen:	dipl, clean, norm, abbr, author_ref, figure, head, hi_rend, lang, lb, lemma, pb, pb_n, persname, pos, reader_ref, script, title

2.25. TheatrumBotanicum_1696_Verzascha

Titel:	Theatrum Botanicum. Das ist: Neu Vollkommenes Kräuter-Buch: Worinnen Allerhand Erdgewächse der Bäumen, Stauden und Kräutern, welche in allen vier Theilen der Welt, sonderlich aber in Europa herfür kommen, neben ihren Fürtrefflichen Würckungen, und deren Gebrauch, wider allerley Kranckheiten an Menschen und Vieh, Mit sonderbahrem Fleiß auff eine ganz neue Art und Weise beschrieben, Auch mit schönen, theils neuen Figuren geziert, und neben denen ordenlichen, so wohl Kräuter- als Kranckheit-Registern, mit nutzlichen Marginalien vorgestellet sind. Allen Aerzten sonderlich auch denen auff dem Land wohnenden höchst nutzlich und ergetzlich
Kürzel:	TheatrumBotanicum_1696_Verzascha
Autor:	Bernhard Verzascha
Jahr:	1696
Ort:	Basel
Verlag:	Bertsch [u.a.]
Seitenangabe:	659-664
Publikationstyp:	Monographie
dipl-Einheiten:	3893
Beschreibung:	Kräutertext
Annotationsebenen:	dipl, clean, norm, abbr, author_ref, figure, head, hi_rend, lang, lb, lemma, note, pb, pb_n, persname, pos, reader_ref, script, title

2.26. SchweizerischerBotanicus_1687_vonRoll

Titel:	Der Schweizerische Botanicus
Kürzel:	SchweizerischerBotanicus_1687_vonRoll
Autor:	Timotheus von Roll
Jahr:	1687
Ort:	NA
Verlag:	Franz Carl Rooß
Seitenangabe:	54-79
Publikationstyp:	Monographie

dipl-Einheiten:	3510
Beschreibung:	Kräutertext
Annotationsebenen:	dipl, clean, norm, abbr, attr_gen, author_ref, comp, comp_orth, head, hi_rend, lang, lb, lemma, morph_ellipsis, pb, pb_n, pos, prot, reader_ref, script, unclear
Anmerkung:	Aus der Originaltranskription der Studierende wurden 2 Seiten gelöscht. Diese warenim der Transkription zugrundeliegenden PDF doppelt vorhanden und wurden von den Studierenden auch doppelt annotiert.

2.27. SonderbaresKraeuterbuch-21-36_1675_Anonymous

Titel:	Curioser Botanicus oder sonderbares Kräuter Buch
Kürzel:	SonderbaresKraeuterbuch-11-36_1675_Anonymous
Autor:	NA
Jahr:	1675
Ort:	NA
Verlag:	NA
Seitenangabe:	21-36
Publikationstyp:	Monographie
dipl-Einheiten:	2876
Beschreibung:	Kräutertext
Annotationsebenen:	dipl, clean, norm, abbr, author_ref, figure, figure_p, head, hi_rend, lang, lb, lemma, pb, pb_n, persname, pos, quote, script, unclear

2.28. SonderbaresKraeuterbuch-11-21_1675_Anonymous

Titel:	Curioser Botanicus oder sonderbares Kräuter Buch
Kürzel:	SonderbaresKraeuterbuch-11-21_1675_Anonymous
Autor:	NA
Jahr:	1675
Ort:	NA
Verlag:	NA
Seitenangabe:	11-21
Publikationstyp:	Monographie
dipl-Einheiten:	1907
Beschreibung:	Kräutertext
Annotationsebenen:	dipl, clean, norm, abbr, attr_gen, comp, comp_orth, figure, figure_p, head, hi_rend, lang, lb, lemma, morph_ellipsis, pb, pb_n, pos, prot, quote, script, unclear

2.29. SonderbaresKraeuterbuch-1-11_1675_Anonymous

Titel:	Curioser Botanicus oder sonderbares Kräuter Buch
Kürzel:	SonderbaresKraeuterbuch-1-11_1675_Anonymous
Autor:	NA
Jahr:	1675
Ort:	NA
Verlag:	NA
Seitenangabe:	1-11
Publikationstyp:	Monographie
dipl-Einheiten:	2245
Beschreibung:	Kräutertext
Annotationsebenen:	dipl, clean, norm, abbr, attr_gen, comp, comp_orth, figure_p, head, hi_rend, lang, lb, lemma, morph_ellipsis, pb, pb_n, pos, prot, quote, script, unclear

2.30. ThesaurusSanitatis_304-321_1673_Nasser

Titel:	Thesaurus Sanitatis
Kürzel:	ThesaurusSanitatis_1673_Nasser
Autor:	Adrian Nasser
Jahr:	1673
Ort:	Nürnberg
Verlag:	Johann Hoffmann
Seitenangabe:	304-321
Publikationstyp:	Monographie
dipl-Einheiten:	4493
Beschreibung:	Kräutertext
Annotationsebenen:	dipl, clean, norm, abbr, author_ref, head, hi_rend, lang, lb, lemma, pb, pb_n, persname, pos, quote, reader_ref, script, title

2.31. ThesaurusSanitatis_1673_Nasser

Titel:	Thesaurus Sanitatis
Kürzel:	ThesaurusSanitatis_1673_Nasser
Autor:	Adrian Nasser
Jahr:	1673
Ort:	Nürnberg
Verlag:	Johann Hoffmann
Seitenangabe:	497-524
Publikationstyp:	Monographie
dipl-Einheiten:	7041
Beschreibung:	Kräutertext
Annotationsebenen:	dipl, clean, norm, abbr, attr_gen, author_ref, comp, comp_orth, head, hi_rend, lang, lb, lemma, morph_ellipsis, pb, pb_n, pos, prot, quote, reader_ref, script, unclear

2.32. Phythologia_1662_Becher

Titel:	Phythologia; Das ist: Deß erläuterten Medicinalischen Parnassi Ander Theil, Nemlich das Kräuter-Buch. Band 2.
Kürzel:	Phythologia_1662_Becher
Autor:	Johann Joachim Becher
Jahr:	1662
Ort:	Ulm
Verlag:	Johann Görlin
Seitenangabe:	305-319
Publikationstyp:	Monographie
dipl-Einheiten:	4321
Beschreibung:	Kräutertext
Annotationsebenen:	dipl, clean, norm, abbr, figure, head, hi_rend, lang, lb, lemma, pb, pb_n, pos, script, title

2.33. Wund-Artzney_1652_Greiff

Titel:	Wund-Artzney
Kürzel:	Wund-Artzney_1652_Greiff
Autor:	Guilelmus Fabricius Hildanus, Friedrich Greiff
Jahr:	1652
Ort:	Frankfurt am Main
Verlag:	Johann Beyer
Seitenangabe:	218-223
Publikationstyp:	Monographie
dipl-Einheiten:	5253
Beschreibung:	Kräutertext
Annotationsebenen:	dipl, clean, norm, abbr, attr_gen, author_ref, head, hi_rend, lang, lb, lemma, morph_ellipsis, note, pb, pb_n, pos, quote, reader_ref, script, unclear, adja_uninfl, comp_amb, comp_lex, comp_n, comp_n_graph, comp_n_mod, comp_a, comp_a_graph, infl_fuge

2.34. PflantzGart-VR_1639_Rhagor

Titel:	Pflantz-Gart (VR)
Kürzel:	PflantzGart-Vorrede_1639_Rhagor
Autor:	Daniel Rhagor
Jahr:	1639
Ort:	Bern
Verlag:	Ben Stephan Schmid
Seitenangabe:	1-10
Publikationstyp:	Monographie
dipl-Einheiten:	2318
Beschreibung:	Kräutertext
Annotationsebenen:	dipl, clean, norm, abbr, attr_gen, author_ref, head, hi_rend, lang, lb, lemma, note, pb, pb_n, pos, quote, reader_ref, script, adja_uninfl, comp_amb, comp_lex, comp_n, comp_n_graph, comp_n_mod, comp_a, comp_a_graph, infl_fuge, kaus, kon, lin, hypo

2.35. PflantzGart-c4_1639_Rhagor

Titel:	Pflantz-Gart (Capitel 4)
Kürzel:	PflantzGart-c4_1639_Rhagor
Autor:	Daniel Rhagor
Jahr:	1639
Ort:	Bern
Verlag:	Ben Stephan Schmid
Seitenangabe:	33-45
Publikationstyp:	Monographie
dipl-Einheiten:	2846
Beschreibung:	Kräutertext
Annotationsebenen:	dipl, clean, norm, abbr, attr_gen, author_ref, head, hi_rend, lang, lb, lemma, morph_ellipsis, note, pb, pb_n, pos, quote, reader_ref, script, unclear, adja_uninfl, comp_amb, comp_lex, comp_n, comp_n_graph, comp_n_mod, comp_a, comp_a_graph, infl_fuge

2.36. PflantzGart_1639_Rhagor

Titel:	Pflantz-Gart
Kürzel:	PflantzGart_1639_Rhagor
Autor:	Daniel Rhagor
Jahr:	1639
Ort:	Bern
Verlag:	Ben Stephan Schmid
Seitenangabe:	92-110
Publikationstyp:	Monographie
dipl-Einheiten:	4067
Beschreibung:	Kräutertext
Annotationsebenen:	dipl, clean, norm, abbr, attr_gen, author_ref, figure, head, hi_rend, lang, lb, lemma, morph_ellipsis, note, pb, pb_n, pos, script, unclear, adja_uninfl, comp_amb, comp_lex, comp_n, comp_n_graph, comp_n_mod, comp_a, comp_a_graph, infl_fuge, kaus, kon, lin, hypo

2.37. Kraeuterbuch_1609_Carrichter

Titel:	Kräutterbuch des Edelen und hochgelehrten herren Bartholomei Carrichters
Kürzel:	Kraeuterbuch_1609_Carrichter
Autor:	Bartholomeus Carrichter
Jahr:	1609
Ort:	Straßburg
Verlag:	Antonium Bertram
Seitenangabe:	47-75
Publikationstyp:	Monographie
dipl-Einheiten:	4992
Beschreibung:	Kräutertext
Annotationsebenen:	dipl, clean, norm, abbr, attr_gen, author_ref, head, hi_rend, lang, lb, lemma, morph_ellipsis, pb, pb_n, pos, reader_ref, script, unclear

2.38 Kraeutterbuch-VR_1609_Carrichter

Titel:	Kräutterbuch des Edelen und hochgelehrten herren Bartholomei Carrichters (Vorrede)
Kürzel:	Kraeuterbuch-VR_1609_Carrichter
Autor:	Bartholomeus Carrichter
Jahr:	1609
Ort:	Straßburg
Verlag:	Antonium Bertram
Seitenangabe:	11-24
Publikationstyp:	Monographie
dipl-Einheiten:	1832
Beschreibung:	Kräutertext
Annotationsebenen:	dipl, clean, norm, abbr, author_ref, column, figure, figure_p, head, hi_rend, lang, lb, lemma, note, pb, pb_n, persname, pos, quote, reader_ref, script, title, unclear

2.39. HortulusSanitatis_1609_Uffenbach

Titel:	Hortulus Sanitatis
Kürzel:	HortulusSanitatis_1609_Uffenbach
Autor:	Peter Uffenbach
Jahr:	1609
Ort:	Frankfurt am Main
Verlag:	Ionae Rhodij
Seitenangabe:	1-21
Publikationstyp:	Monographie
dipl-Einheiten:	6516
Beschreibung:	Kräutertext
Annotationsebenen:	dipl, clean, norm, abbr, attr_gen, author_ref, head, hi_rend, lang, lb, lemma, morph_ellipsis, note, pb, pb_n, pos, quote, script, unclear, adja_uninfl, comp_amb, comp_lex, comp_n, comp_n_graph, comp_n_mod, comp_a, comp_a_graph, infl_fuge

2.40. HortulusSanitatis-VR_1609_Uffenbach

Titel:	Hortulus Sanitatis (Vorrede)
Kürzel:	HortulusSanitatis-VR_1609_Uffenbach
Autor:	Peter Uffenbach
Jahr:	1609
Ort:	Frankfurt am Main
Verlag:	Ionae Rhodij
Seitenangabe:	NA
Publikationstyp:	Monographie
dipl-Einheiten:	3379
Beschreibung:	Kräutertext
Annotationsebenen:	dipl, clean, norm, abbr, author_ref, column, figure, figure_p, head, hi_rend, lang, lb, lemma, note, pb, pb_n, persname, pos, quote, quotemark, reader_ref, script, title, unclear

2.41. AlchymistischePractic_1603_Libavius

Titel:	Alchymistische Practic
Kürzel:	AlchymistischePractic_1603_Libavius
Autor:	Andreas Libavius
Jahr:	1603
Ort:	Frankfurt
Verlag:	Johann Saur
Seitenangabe:	5-26
Publikationstyp:	Monographie
dipl-Einheiten:	5063
Beschreibung:	Kräutertext
Annotationsebenen:	dipl, clean, norm, abbr, attr_gen, author_ref, head, hi_rend, lang, lb, lemma, morph_ellipsis, note, pb, pb_n, pos, quote, reader_ref, script, unclear, adja_uninfl, comp_amb, comp_lex, comp_n, comp_n_graph, comp_n_mod, comp_a, comp_a_graph, infl_fuge

2.42. AlchymistischePractic-VR_1603_Libavius

Titel:	Alchymistische Practic (Vorrede)
Kürzel:	AlchymistischePractic-VR_1603_Libavius
Autor:	Andreas Libavius
Jahr:	1603
Ort:	Frankfurt
Verlag:	Johann Saur
Seitenangabe:	3-4
Publikationstyp:	Monographie
dipl-Einheiten:	441
Beschreibung:	Kräutertext
Annotationsebenen:	dipl, clean, norm, abbr, author_ref, head, lang, lb, lemma, pb, pb_n, persname, pos, quote, reader_ref, script, title, adja_uninfl, comp_amb, comp_lex, comp_n, comp_n_graph, comp_n_mod, comp_a, comp_a_graph, infl_fuge

2.43. NeuwKreuterbuch_1588_Theodorus

Titel:	Neuw Kreuterbuch
Kürzel:	NeuwKreuterbuch_1588_Theodorus
Autor:	Tabernaemontanus, Jacobus Theodorus
Jahr:	1588
Ort:	Frankfurt am Main
Verlag:	NA
Seitenangabe:	33-44
Puplikationstyp:	Monographie
dipl-Einheiten:	9761
Beschreibung:	Kräutertext
Annotationsebene:	dipl, clean, norm, abbr, author_ref, column, figure, figure_p, head, hi_rend, lang, lb, lemma, note, pb, pb_n, persname, pos, reader_ref, script, title, unclear

2.44. NeuwKreuterbuch-VR_1588_Theodorus

Titel:	Neuw Kreuterbuch (Vorrede)
Kürzel:	NeuwKreuterbuch-VR_1588_Theodorus
Autor:	Tabernaemontanus, Jacobus Theodorus
Jahr:	1588
Ort:	Frankfurt am Main
Verlag:	NA
Seitenangabe:	NA
Puplikationstyp:	Monographie
dipl-Einheiten:	6656
Beschreibung:	Kräutertext
Annotationsebene:	dipl, clean, norm, abbr, author_ref, column, figure, figure_p, head, hi_rend, lang, lb, lemma, note, pb, pb_n, persname, pos, quote, reader_ref, script, title, unclear

2.45. Paradeiszgaertlein_1588_Rosbach

Titel:	Paradeiszgärtlein
Kürzel:	Paradeiszgaertlein_1588_Rosbach
Autor:	Conrad Rosbach
Jahr:	1588
Ort:	Frankfurt am Main
Verlag:	Johann Spieß
Seitenangabe:	1-43
Publikationstyp:	Monographie
dipl-Einheiten:	5054
Beschreibung:	Kräutertext
Annotationsebenen:	dipl, clean, norm, abbr, attr_gen, author_ref, comp, comp_orth, figure, head, hi_rend, lang, lb, lemma, morph_ellipsis, pb, pb_n, pos, prot, quote, reader_ref, reference, script, unclear, adja_uninfl, comp_amb, comp_lex, comp_n, comp_n_graph, comp_n_mod, comp_a, comp_a_graph, infl_fuge

2.46. Paradeiszgaertlein-VR_1588_Rosbach

Titel:	Paradeiszgärtlein (Vorrede)
Kürzel:	Paradeiszgaertlein-VR_1588_Rosbach
Autor:	Conrad Rosbach
Jahr:	1588
Ort:	Frankfurt am Main
Verlag:	Johann Spieß
Seitenangabe:	NA
Publikationstyp:	Monographie
dipl-Einheiten:	2678
Beschreibung:	Kräutertext
Annotationsebenen:	dipl, clean, norm, abbr, author_ref, column, figure, figure_p, head, hi_rend, lang, lb, lemma, note, pb, pb_n, persname, pos, quote, reader_ref, script, title, unclear

2.47. NewKreueterbuch_1563_Handsch

Titel:	New Kreüterbuch: Mit den allerschönsten vnd artlichsten Figuren aller Gewechß, dergleichen vormals in keiner sprach nie an tag kommen
Kürzel:	NewKreueterbuch_1563_Handsch
Autor:	Georg Handsch
Jahr:	1563
Ort:	Prag
Verlag:	Melantrich von Auentin und Valgriß
Seitenangabe:	277-283
Publikationstyp:	Monographie
dipl-Einheiten:	3785
Beschreibung:	Kräutertext
Annotationsebenen:	dipl, clean, norm, abbr, author_ref, figure, head, hi_rend, lang, lb, lemma, pb, pb_n, persname, pos, reader_ref, script, adja_uninfl, comp_amb, comp_lex, comp_n, comp_n_graph, comp_n_mod, comp_a, comp_a_graph, infl_fuge

2.48. NewKreuterbuch-VR_1563_Handsch

Titel:	New Kreüterbuch: Mit den allerschönsten vnd artlichsten Figuren aller Gewechß, dergleichen vormals in keiner sprach nie an tag kommen (Vorrede)
Kürzel:	NewKreueterbuch-VR_1563_Handsch
Autor:	Georg Handsch
Jahr:	1563
Ort:	Prag
Verlag:	Melantrich von Auentin und Valgriß
Seitenangabe:	NA
Publikationstyp:	Monographie
dipl-Einheiten:	2662
Beschreibung:	Kräutertext
Annotationsebenen:	dipl, clean, norm, abbr, author_ref, column, figure, figure_p, head, hi_rend, lang, lb, lemma, note, pb, pb_n, persname, pos, quote, reader_ref, script, title, unclear

2.49. WieSichMeniglich_1557_vonBodenstein

Titel:	Wie sich meniglich von dem Cyperlin, so Podagra genennet, hüten soll
Kürzel:	WieSichMeniglich_1557_vonBodenstein
Autor:	von Bodenstein, Adam
Jahr:	1557
Ort:	Basel
Verlag:	Bartholome Stähälin
Seitenangabe:	28-47
Publikationstyp:	Monographie
dipl-Einheiten:	5508
Beschreibung:	Kräutertext
Annotationsebenen:	dipl, clean, norm, abbr, attr_gen, author_ref, head, lang, lb, lemma, morph_ellipsis, pb, pb_n, persname, pos, quote, reader_ref, script, unclear, adja_uninfl, comp_amb, comp_lex, comp_n, comp_n_graph, comp_n_mod, comp_a, comp_a_graph, infl_fuge

2.50. WieSichMeniglich-VR_1557_vonBodenstein

Titel:	Wie sich meniglich von dem Cyperlin, so Podagra genennet, hüten soll (Vorrede)
Kürzel:	WieSichMeniglich-VR_1557_vonBodenstein
Autor:	von Bodenstein, Adam
Jahr:	1557
Ort:	Basel
Verlag:	Bartholome Stähälin
Seitenangabe:	NA
Publikationstyp:	Monographie
dipl-Einheiten:	707
Beschreibung:	Kräutertext
Annotationsebenen:	dipl, clean, norm, abbr, author_ref, head, hi_rend, lang, lb, lemma, pb, persname, pos, quote, reader_ref, script, adja_uninfl, comp_amb, comp_lex, comp_n, comp_n_graph, comp_n_mod, comp_a, comp_a_graph, infl_fuge

2.51. Experimenta_1550_Schellenberg

Titel:	Experimenta von zwentzig Pestilentz Wurtzelen unnd Kreutter
Kürzel:	Experimenta_1550_Schellenberg
Autor:	Schnellenberg, Tarquinius
Jahr:	1550
Ort:	Strassburg
Verlag:	Rihel
Seitenangabe:	33-43
Publikationstyp:	Monographie
dipl-Einheiten:	5635
Beschreibung:	Kräutertext
Annotationsebenen:	dipl, clean, norm, abbr, author_ref, column, figure, figure_p, head, hi_rend, lang, lb, lemma, note, pb, pb_n, persname, pos, reader_ref, script, title, unclear

2.52. Experimenta-VR_1550_Schellenbeg

Titel:	Experimenta von zwentzig Pestilentz Wurtzelen unnd Kreutter (Vorrede)
Kürzel:	Experimenta-VR_1550_Schellenberg
Autor:	Schnellenberg, Tarquinius
Jahr:	1550
Ort:	Strassburg
Verlag:	Rihel
Seitenangabe:	2-3
Publikationstyp:	Monographie
dipl-Einheiten:	491
Beschreibung:	Kräutertext
Annotationsebenen:	dipl, clean, norm, abbr, author_ref, column, figure, figure_p, head, hi_rend, lang, lb, lemma, note, pb, pb_n, persname, pos, quote, reader_ref, script, title, unclear

2.53. NewKreueterbuch_1543_Fuchs

Titel:	New Kreüterbuch
Kürzel:	NewKreueterbuch_1543_Fuchs
Autor:	Fuchs, Leonhard
Jahr:	1543
Ort:	Basel
Verlag:	Michael Isingrin
Seitenangabe:	2-e4
Publikationstyp:	Monographie
dipl-Einheiten:	5221
Beschreibung:	Kräutertext
Annotationsebenen:	dipl, clean, norm, abbr, attr_gen, author_ref, figure, head, hi_rend, lang, lb, lemma, morph_ellipsis, pb, pb_n, persname, pos, quote, reader_ref, script, adja_uninfl, comp_amb, comp_lex, comp_n, comp_n_graph, comp_n_mod, comp_a, comp_a_graph, infl_fuge

2.54. NewKreuterbuch-c19-25_1543_Fuchs

Titel:	New Kreüterbuch
Kürzel:	NewKreuterbuch-c19-25_1543_Fuchs
Autor:	Fuchs, Leonhard
Jahr:	1543
Ort:	Basel
Verlag:	Michael Isingrin
Seitenangabe:	Kapitel 19-25
Publikationstyp:	Monographie
dipl-Einheiten:	2821
Beschreibung:	Kräutertext
Annotationsebenen:	dipl, clean, norm, abbr, author_ref, figure, figure_p, head, hi_rend, lang, lb, lemma, note, pb, pos, quote, persname, reader_ref, script, title, comp_n, comp_n_graph, comp_n_mod, infl_fuge, attr_gen, adja_uninfl, morph_ellipsis, comp_a, comp_a_graph, comp_lex, comp_amb

2.55. NewKreuterbuch-VR_1543_Fuchs

Titel:	New Kreüterbuch (Vorrede)
Kürzel:	NewKreuterbuch-VR_1543_Fuchs
Autor:	Fuchs, Leonhard
Jahr:	1543
Ort:	Basel
Verlag:	Michael Isingrin
Seitenangabe:	NA
Publikationstyp:	Monographie
dipl-Einheiten:	2221
Beschreibung:	Kräutertext
Annotationsebenen:	dipl, clean, norm, abbr, author_ref, column, figure, figure_p, head, hi_rend, lang, lb, lemma, note, pb, pb_n, persname, pos, quote, reader_ref, script, title, unclear

2.56. NewKreuetterBuch_1539_Bock

Titel:	New Kreütter Buch: von Underscheydt, Würckung und Namen der Kreütter so in teütschen Landen wachsen; auch der selbigen eygentlichem und wolgegründtem Gebrauch in der Artznei, zu behalten und zu fürdern Leibs Gesuntheyt fast nutz und tröstlichen, vorab gemeynem Verstand
Kürzel:	NewKreuetterBuch_1539_Bock
Autor:	Hieronymus Bock
Jahr:	1539
Ort:	Straßburg
Verlag:	Rihel
Seitenangabe:	XCVI-C
Publikationstyp:	Monographie
dipl-Einheiten:	4408
Beschreibung:	Kräutertext
Annotationsebenen:	dipl, clean, norm, abbr, author_ref, head, hi_rend, lang, lb, lemma, note, pb, pb_n, persname, pos, reader_ref, script, title, unclear, adja_uninfl, comp_amb, comp_lex, comp_n, comp_n_graph, comp_n_mod, comp_a, comp_a_graph, infl_fuge

2.57. NewKreuetterbuch-CXXXIIIJ-CXXXVIIJ_1539_Bock

Titel:	New Kreütter Buch: von Underscheydt, Würckung und Namen der Kreütter so in teütschen Landen wachsen; auch der selbigen eygentlichem und wolgegründtem Gebrauch in der Artznei, zu behalten und zu fürdern Leibs Gesuntheyt fast nutz und tröstlichen, vorab gemeynem Verstand
Kürzel:	NewKreuetterbuch-CXXXIIIJ-CXXXVIIJ_1539_Bock
Autor:	Hieronymus Bock
Jahr:	1539
Ort:	Straßburg
Verlag:	Rihel
Seitenangabe:	CXXXIIIJ-CXXXVIIJ
Publikationstyp:	Monographie
dipl-Einheiten:	4055
Beschreibung:	Kräutertext
Annotationsebenen:	dipl, clean, norm, abbr, author_ref, head, hi_rend, lang, lb, lemma,

	note, pb, pb_n, persname, pos, reader_ref, script, title, comp_n, comp_n_graph, comp_n_mod, infl_fuge, attr_gen, adja_uninfl, morph_ellipsis, comp_a, comp_a_graph, comp_lex, comp_amb

2.58. NewKreuetterBuch-VR_1539_Bock

Titel:	New Kreütter Buch: von Underscheydt, Würckung und Namen der Kreütter so in teütschen Landen wachsen; auch der selbigen eygentlichem und wolgegründtem Gebrauch in der Artznei, zu behalten und zu fürdern Leibs Gesuntheyt fast nutz und tröstlichen, vorab gemeynem Verstand (Vorrede)
Kürzel:	NewKreuetterBuch_1539_Bock
Autor:	Hieronymus Bock
Jahr:	1539
Ort:	Straßburg
Verlag:	Rihel
Seitenangabe:	VII-X
Publikationstyp:	Monographie
dipl-Einheiten:	2021
Beschreibung:	Kräutertext
Annotationsebenen:	dipl, clean, norm, abbr, author_ref, figure, head, hi_rend, lang, lb, lemma, note, pb, pb_n, persname, pos, quote, reader_ref, script, title, unclear, adja_uninfl, comp_amb, comp_lex, comp_n, comp_n_graph, comp_n_mod, comp_a, comp_a_graph, infl_fuge

2.59. ContrafaytKreuterbuch_1532_Brunfels

Titel:	Contrafayt kreüterbuch
Kürzel:	ContrafaytKreuterbuch_1532_Brunfels
Autor:	Otto Brunfels
Jahr:	1532
Ort:	Straßburg
Verlag:	Hans Schotten
Seitenangabe:	312-323

Publikationstyp:	Monographie
dipl-Einheiten:	4359
Beschreibung:	Kräutertext
Annotationsebenen:	dipl, clean, norm, abbr, attr_gen, author_ref, figure, head, hi_rend, lang, lb, lemma, pb, pb_n, persname, pos, reader_ref, script, unclear, adja_uninfl, comp_amb, comp_lex, comp_n, comp_n_graph, comp_n_mod, comp_a, comp_a_graph, infl_fuge

2.60. ContrafaytKreuterbuch- CCXXXVII-CCXLVIII_1532_Brunfels

Titel:	Contrafayt kreüterbuch
Kürzel:	ContrafaytKreuterbuch- CCXXXVII-CCXLVIII_1532_Brunfels
Autor:	Otto Brunfels
Jahr:	1532
Ort:	Straßburg
Verlag:	Hans Schotten
Seitenangabe:	CCXXXVII-CCXLVIII
Publikationstyp:	Monographie
dipl-Einheiten:	3988
Beschreibung:	Kräutertext
Annotationsebenen:	dipl, clean, norm, abbr, author_ref, figure, hi_rend, lang, lb, lemma, pb, pb_n, persname, pos, quote, reader_ref, script, adja_uninfl, comp_amb, comp_lex, comp_n, comp_n_graph, comp_n_mod, comp_a, comp_a_graph, infl_fuge

2.61. ContrafaytKreuterbuch-CCLVII-CCLXX_1532_Brunfels

Titel:	Contrafayt kreüterbuch
Kürzel:	ContrafaytKreuterbuch-CCLVII-CCLXX_1532_Brunfels
Autor:	Otto Brunfels
Jahr:	1532

Ort:	Straßburg
Verlag:	Hans Schotten
Seitenangabe:	CCLVII-CCLXX
Publikationstyp:	Monographie
dipl-Einheiten:	3975
Beschreibung:	Kräutertext
Annotationsebenen:	dipl, clean, norm, abbr, author_ref, figure, figure_p, hi_rend, lang, lb, lemma, note, pb, pb_n, persname, pos, quote, reader_ref, script, title, comp_n, comp_n_graph, comp_n_mod, infl_fuge, attr_gen, adja_uninfl, morph_ellipsis, comp_a, comp_a_graph, comp_lex, comp_amb

2.62. ContrafaytKreuterbuch-CCLXXI-CCXC_1532_Brunfels

Titel:	Contrafayt kreüterbuch
Kürzel:	ContrafaytKreuterbuch-CCLXXI-CCXC_1532_Brunfels
Autor:	Otto Brunfels
Jahr:	1532
Ort:	Straßburg
Verlag:	Hans Schotten
Seitenangabe:	CCLXXI-CCXC
Publikationstyp:	Monographie
dipl-Einheiten:	4401
Beschreibung:	Kräutertext
Annotationsebenen:	dipl, clean, norm, abbr, author_ref, figure, figure_p, hi_rend, lang, lb, lemma, note, pb, pb_n, persname, pos, quote, reader_ref, script, title, comp_n, comp_n_graph, comp_n_mod, infl_fuge, attr_gen, adja_uninfl, morph_ellipsis, comp_a, comp_a_graph, comp_lex, comp_amb

2.63. ContrafaytKreuterbuch-VR_1532_Brunfels

Titel:	Contrafayt kreüterbuch (Vorrede)
Kürzel:	ContrafaytKreuterbuch-VR_1532_Brunfels
Autor:	Otto Brunfels
Jahr:	1532
Ort:	Straßburg
Verlag:	Hans Schotten
Seitenangabe:	VIII-IX
Publikationstyp:	Monographie
dipl-Einheiten:	664
Beschreibung:	Kräutertext
Annotationsebenen:	dipl, clean, norm, abbr, author_ref, figure, head, lang, lb, lemma, note, pb, pb_n, persname, pos, quote, script, title, adja_uninfl, comp_amb, comp_lex, comp_n, comp_n_graph, comp_n_mod, comp_a, comp_a_graph, infl_fuge

2.64. ArtzneyBuchleinDerKreutter_1532_Tallat

Titel:	Artzney Buchlein der kreutter
Kürzel:	ArtzneyBuchleinDerKreutter_1532_Tallat
Autor:	Johannes Tallat
Jahr:	1532
Ort:	Leipzig
Verlag:	Michael Blum
Seitenangabe:	NA
Publikationstyp:	Monographie
dipl-Einheiten:	4630
Beschreibung:	Kräutertext
Annotationsebenen:	dipl, clean, norm, abbr, attr_gen, author_ref, head, hi_rend, lang, lb, lemma, note, pb, persname, pos, reader_ref, script, title, unclear, adja_uninfl, comp_amb, comp_lex, comp_n, comp_n_graph, comp_n_mod, comp_a, comp_a_graph, infl_fuge

2.65. ArtzneyBuchleinDerKreutter-Abrotanum_1532_Tallat

Titel:	Artzney Buchlein der kreutter
Kürzel:	ArtzneyBuchleinDerKreutter-Abrotanum_1532_Tallat
Autor:	Johannes Tallat
Jahr:	1532
Ort:	Leipzig
Verlag:	Michael Blum
Seitenangabe:	ab Abrotanum
Publikationstyp:	Monographie
dipl-Einheiten:	3530
Beschreibung:	Kräutertext
Annotationsebenen:	dipl, clean, norm, abbr, author_ref, head, hi_rend, lang, lb, lemma, pb, persname, pos, quote, reader_ref, script, title, unclear, comp_n, comp_n_graph, comp_n_mod, infl_fuge, attr_gen, adja_uninfl, morph_ellipsis, comp_a, comp_a_graph, comp_lex, comp_amb, kaus, kon, lin, hypo

2.66. ArtzneyBuchleinDerKreutter-Cretanus_1532_Tallat

Titel:	Artzney Buchlein der kreutter
Kürzel:	ArtzneyBuchleinDerKreutter-Cretanus_1532_Tallat
Autor:	Johannes Tallat
Jahr:	1532
Ort:	Leipzig
Verlag:	Michael Blum
Seitenangabe:	Cretanus bis Dens leonis
Publikationstyp:	Monographie
dipl-Einheiten:	1120
Beschreibung:	Kräutertext
Annotationsebenen:	dipl, clean, norm, infl_fuge, comp_amb, comp_n, comp_n_graph, comp_n_mod, comp_lex, comp_a, comp_a_graph, adja_uninfl, attr_gen, morph_ellipsis, lang, abbr, pb, pb_n, lb, head, hi_rend, script, author_ref, quote

2.67. ArtzneyBuchleinDerKreutter-VR_1532_Tallat

Titel:	Artzney Buchlein der kreutter (Vorrede)
Kürzel:	ArtzneyBuchleinDerKreutter-VR_1532_Tallat
Autor:	Johannes Tallat
Jahr:	1532
Ort:	Leipzig
Verlag:	Michael Blum
Seitenangabe:	NA
Publikationstyp:	Monographie
dipl-Einheiten:	270
Beschreibung:	Kräutertext
Annotationsebenen:	dipl, clean, norm, abbr, author_ref, figure, head, hi_rend, lang, lb, lemma, pb, persname, pos, quote, reader_ref, script, adja_uninfl, comp_amb, comp_lex, comp_n, comp_n_graph, comp_n_mod, comp_a, comp_a_graph, infl_fuge, kaus, kon, lin, hypo

2.68. GartDerGesundheit_1487_vonCuba

Titel:	Gart der Gesundheit
Kürzel:	GartDerGesundheit_1487_vonCuba
Autor:	Johannes von Cuba
Jahr:	1487
Ort:	Ulm
Verlag:	NA
Seitenangabe:	NA
Publikationstyp:	Monographie
dipl-Einheiten:	4700
Beschreibung:	Kräutertext
Annotationsebenen:	dipl, clean, norm, abbr, attr_gen, author_ref, clause_type, figure, head, hi_rend, lang, lb, lemma, pb, persname, pos, reader_ref, script, title, unclear, adja_uninfl, comp_amb, comp_lex, comp_n, comp_n_graph, comp_n_mod, comp_a, comp_a_graph, infl_fuge

2.69. GartDerGesundheit-c15-23_1487_Cuba

Titel:	Gart der Gesundheit
Kürzel:	GartDerGesundheit_1487_vonCuba
Autor:	Johannes von Cuba
Jahr:	1487
Ort:	Ulm
Verlag:	NA
Seitenangabe:	Kapitel 15-23
Publikationstyp:	Monographie
dipl-Einheiten:	3922
Beschreibung:	Kräutertext
Annotationsebenen:	dipl, clean, norm, abbr, author_ref, column, figure, head, hi_rend, lang, lb, lemma, pb, persname, pos, quote, reader_ref, script, title, unclear, comp_n, comp_n_graph, comp_n_mod, infl_fuge, attr_gen, adja_uninfl, morph_ellipsis, comp_a, comp_a_graph, comp_lex, comp_amb

2.70. GartDerGesundheit-c23-33_1487_Cuba

Titel:	Gart der Gesundheit
Kürzel:	GartDerGesundheit-c23-33_1487_Cuba
Autor:	Johannes von Cuba
Jahr:	1487
Ort:	Ulm
Verlag:	NA
Seitenangabe:	Kapitel 23-33
Publikationstyp:	Monographie
dipl-Einheiten:	3843
Beschreibung:	Kräutertext
Annotationsebenen:	dipl, clean, norm, abbr, author_ref, column, figure, head, hi_rend, lang, lb, lemma, pb, persname, pos, quote, reader_ref, script, title, comp_n, comp_n_graph, comp_n_mod, infl_fuge, attr_gen, adja_uninfl, comp_lex, comp_amb, kaus, kon, lin, hypo

2.71. GartDerGesundheit-VR_1487_vonCuba

Titel:	Gart der Gesundheit (Vorrede)
Kürzel:	GartDerGesundheit-VR_1487_vonCuba
Autor:	Johannes von Cuba
Jahr:	1487
Ort:	Ulm
Verlag:	NA
Seitenangabe:	aij-aiij
Publikationstyp:	Monographie
dipl-Einheiten:	1352
Beschreibung:	Kräutertext
Annotationsebenen:	dipl, clean, norm, abbr, author_ref, head, lang, lb, lemma, pb, pb_n, persname, pos, reader_ref, script, title, unclear, adja_uninfl, comp_amb, comp_lex, comp_n, comp_n_graph, comp_n_mod, comp_a, comp_a_graph, infl_fuge, kaus, kon, lin, hypo

2.72. BuchDerNatur_1482_vonMegenberg

Titel:	Das Buch der Natur
Kürzel:	BuchDerNatur_1482_vonMegenberg
Autor:	Conradus von Megenberg
Jahr:	1482
Ort:	Augsburg
Verlag:	NA
Seitenangabe:	NA
Publikationstyp:	Monographie
dipl-Einheiten:	5215
Beschreibung:	Kräutertext
Annotationsebenen:	dipl, clean, norm, abbr, author_ref, head, hi_rend, lang, lb, lemma, pb, persname, pos, reader_ref, script, adja_uninfl, comp_amb, comp_lex, comp_n, comp_n_graph, comp_n_mod, comp_a, comp_a_graph, infl_fuge

2.73. BuchDerNatur-Canamellis_1482_vonMegenberg

Titel:	Das Buch der Natur
Kürzel:	BuchDerNatur-Canamellis_1482_vonMegenberg
Autor:	Conradus von Megenberg
Jahr:	1482
Ort:	Augsburg
Verlag:	NA
Seitenangabe:	NA
Publikationstyp:	Monographie
dipl-Einheiten:	5077
Beschreibung:	Kräutertext
Annotationsebenen:	dipl, clean, norm, abbr, author_ref, column, figure, figure_p, head, hi_rend, lang, lb, lemma, note, pb, pb_n, persname, pos, quote, reader_ref, script, title, unclear

3. Annotationsebenen – Transkription/Normalisierung

3.1. dipl

Typ: *Layer* - dipl

Name:	dipl
Dokumentation:	Siehe Ridges-Website: https://www.linguistik.hu-berlin.de/de/institut/professuren/korpuslinguistik/forschung/ridges-projekt/documentation Beispiele: Siehe Anhang
dipl-Einheiten:	305056
Beschreibung:	Die Transkription von Faksimiles stellt für die korpuslinguistische Aufbereitung zumeist die grundlegende, diplomatische Ebene (**dipl**). Diese Ebene soll sich graphisch hinsichtlich Orthographie, Getrennt- und Zusammenschreibung und Sonderzeichen möglichst nah am zugrunde liegenden Faksimile orientieren. Grundsätzlich wird so auch entgegen modernen Orthographieregeln segmentiert oder transkribiert. ȝ wird nicht von *Z* unterschieden. *ß* wird in Frakturtexten als *ſz* transkribiert, da hier eine Unterscheidung zwischen der Ligatur und den beiden Einzelzeichen häufig schwierig ist.

Typ: Preparationstep – dipl

Die dipl-Ebene wird für gewöhnlich folgendermaßen erstellt (für die Aufbereitungsschritte mit ocr in den Dokumenten der Version 7 siehe hier):

Schritt:	1a
Aufbereitung:	Transkript
Aufbereitungsart:	manuell
Tool:	NA
Format:	Text-Datei
Segmentierung der Annotation:	eigenständig
Qualitätsprüfer:	Korpling
Datum:	2019
Qualitätsprüfung:	manuell
Editor:	Studierende der Humboldt-Universität zu Berlin und der Georg-August-Universität Göttingen

Schritt:	2a
Aufbereitung:	Tokenisierung
Aufbereitungsart:	Automatisch
Tool:	Treetagger Version 3.2
Format:	Treetagger-Output
Segmentierung der Annotation:	eigenständig
Qualitätsprüfer:	Studierende der Humboldt-Universität zu Berlin und der Georg-August-Universität Göttingen, korpling
Datum:	2019
Qualitätsprüfung:	manuelle Konsistenzprüfung
Editor:	Catharina Fischer

Schritt:	3a
Aufbereitung:	Konvertierung von Excel 2013 nach PAULA-XML und ANNIS

Aufbereitungsart:	automatisch
Tool:	Pepper, Konvertierung des Formats Excel nach ANNIS und PAULA mithilfe der folgenden Importer: SpreadsheetImporter. Export ins ANNIS- und PAULA-Format mit dem ANNISExporter und dem PAULAExporter (https://korpling.german.hu-berlin.de/saltnpepper/pepper/download/stable/Pepper_2019.06.11.zip)
Format:	PAULA-XML, ANNIS
Segmentierung der Annotation:	eigenständig
Qualitätsprüfer:	korpling
Datum:	2019
Qualitätsprüfung:	NA
Editor:	Korpling

3.2. clean

Typ: *Layer* – clean

Name:	clean
Dokumentation:	Siehe Ridges-Website http://korpling.org/ridges Beispiele: Siehe Anhang
clean-Einheiten:	297091
Beschreibung:	Die clean-Ebene enthält erste vollautomatisch erstellte Normalisierungen hinsichtlich Sonderzeichen und graphischer Strukturierungen. So werden u.a. Ligaturen, die nach moderner Rechtschreibung nicht mehr verwendet werden, normalisiert. Graphische Markierungen der einzelnen Texte wie Zeilenumbrüche werden aufgelöst und Sonderzeichen einiger Fonts wie Fraktur auf die heutigen Schriftsätze übertragen. Die Dokumente, die seit der Version 4.0 dazugekommen sind, stellen neue Anforderungen an die automatisch erstellte clean-Ebene. Mit Tilden versehene Vokale können nicht länger kontextunabhängig normalisiert werden. Deshalb werden ab Version 4.0 für Token, die Vokale mit Tilden enthalten, alle möglichen Formen dieser Token in der clean angegeben. Die verschiedenen Formen werden durch ‚|' getrennt (zum Beispiel: *aufzwēdig* wird zu *auszwemdig*|*auszwendig*). In der clean-Ebene werden außerdem Wörter, die durch einen Zeilenumbruch getrennt und mit einem Bindeelement versehen sind, zusammengezogen. Beginnt das zweite Wort mit einem Großbuchstaben, wird dieser in der clean-Ebene in Kleinschreibung

realisiert. Ist der komplette zweite Bestandteil in Großbuchstaben geschrieben, bleibt dies so bestehen (*Gelb-Sucht* wird zu *Gelbsucht*; *MON-* <lb> *TANUM* wird zu *MONTANUM*).

Trunkierte Elemente, die am Zeilenende stehen, werden bisher nicht als solche erkannt und daher automatisch mit dem ersten Element der folgenden Zeile zusammengezogen (*Speifz-und Nahrungs-Saffts* wird zu *Speiszund* (sic!)*Nahrungssaffts*).

Für eine komplette Auflistung aller Ersetzungen, die für die Erstellung der clean-Ebene gemacht wurden, siehe die Readme zum Skript, das zur Erstellung benutzt wurde.

Typ: Preparationstep – clean

Schritt:	1
Aufbereitung:	Normalisierung, Annotation
Aufbereitungsart:	Automatisch
Tool	clean-skript_V3.py (https://hu.berlin/clean-script-v3)
Format:	Clean-Skript-Output
Segmentierung der Annotation:	eigenständig
Qualitätsprüfer:	Korpling
Datum:	2019
Qualitätsprüfung:	manuelle Konsistenzprüfung
Editor:	Korpling

Schritt:	2
Aufbereitung:	Konvertierung von Excel 2013 nach PAULA-XML und ANNIS
Aufbereitungsart:	Automatisch
Tool:	Pepper, Konvertierung des Formats Excel nach ANNIS und PAULA mithilfe der folgenden Importer: SpreadsheetImporter. Export ins ANNIS- und PAULA-Format mit dem ANNISExporter und dem PAULAExporter (https://korpling.german.hu-berlin.de/saltnpepper/pepper/download/stable/Pepper_2019.06.11.zip)
Format:	PAULA-XML, ANNIS
Segmentierung der Annotation:	Eigenständig

Qualitätsprüfer:	Korpling
Datum:	2019
Qualitätsprüfung:	NA
Editor:	Korpling

3.3. norm

Typ: *Layer* – norm

Name:	Norm
Dokumentation:	Siehe Ridges-Website http://korpling.org/ridges Beispiele: Siehe Anhang
norm-Einheiten:	294955
Beschreibung:	Die norm-Ebene beinhaltet normalisierte Annotationen der clean-Token. Die Normalisierung erfolgt hinsichtlich Tokenisierung, Graphematik, Phonologie und Morphologie. Normalisierungsideal ist die gegenwartssprachliche deutsche Norm (standarddt.), die durch das Duden-Wörterbuch und die Duden-Grammatik repräsentiert wird. **Tokenisierung**: Die Norm-Ebene repräsentiert die standarddt. lexikalischen Einheiten: Rinder Zunge > Rinderzunge **Graphematik**: Die historische graphematische Repräsentation wird an die aktuelle orthographische Norm angepasst. kreutter → Kräuter Zur Normalisierung von Interpunktionszeichen siehe Anhang. **Phonologie**: Diachrone und dialektale Varianten werden dem nhd. phonologischen System angepasst: wänlîn → Wännlein lehret → lehrt

Morphologie
Fehlende oder abweichende Morpheme werden hinzugefügt oder durch im NHD grammatische Morpheme ersetzt:
worden -> geworden
das kal Haupt -> das kahle Haupt
Paradigmatische Abweichungen werden ebenfalls normalisiert:
Mannen -> Männern
bu -> baute

Nicht normalisiert wird:
Morphosyntax:
Im standarddt. prinzipiell grammatische Morpheme werden nicht normalisiert, auch wenn sie morphosyntaktisch
im Standarddt. durch ein anderes Morphem ersetzt werden müssten.
die verserte Därmelein -> die versehrte Därmelein (nicht: die versehrten Därmelein)

Syntax:
Die Positionen der Konstituenten werden nicht normalisiert.
Aus standarddt. Sicht ungrammatische Elemente werden nicht entfernt.
Aus standarddt. Sicht fehlende Konstituenten (z.B. Artikel) werden nicht eingefügt.
Aus standarddt. Sicht ungrammatische Kongruenzabweichungen werden nicht normalisiert.
Aus standarddt. Sicht ungrammatische Kasusrektion wird nicht normalisiert.

Lexik:
Aus standarddt. Sicht nicht mehr oder anders gebrauchte Lexeme werden nicht durch die standarddt.Entsprechung
normalisiert.
Hornung -> Hornung (nicht:September)
Die Existenz eines standarddeutschen Lexems wird durch die Lemmatisierung im Duden definiert.

Die Existenz (und graphematische Repräsentation) ausgestorbener Lexeme wird durch die Lemmatisierung
in Grimms Wörterbuch vorgegeben.
Zur Nutzung von Nachschlagewerken für die Erstellung der norm-Ebene siehe XXX
Begründete Einzelentscheidungen siehe Anhang.

Pragmatik:
Aus standarddt. Sich ungewöhnlich realisierte kommunikative Funktionen (z.B. Anreden, Ehrbezeugungen,
wörtliche Zitationskennzeichnungen) werden nicht normalisiert.

Beispiele und Einzelentscheidungen siehe Anhang.

Typ: Preparationstep – norm

Schritt:	1
Aufbereitung:	Normalisierung, Annotation
Aufbereitungsart:	Manuell
Tool:	Excel 2013
Format:	Excel 2013
Segmentierung der Annotation:	Eigenständig
Qualitätsprüfer:	Korpling
Datum:	2019
Qualitätsprüfung:	manuelle Korrektur
Editor	Studierende der Humboldt-Universität zu Berlin und der Georg-August-Universität Göttingen

Schritt:	2
Aufbereitung:	Konvertierung von Excel 2013 nach PAULA-XML und ANNIS
Aufbereitungsart:	Automatisch
Tool:	Pepper, Konvertierung des Formats Excel nach ANNIS und PAULA mithilfe der folgenden Importer: SpreadsheetImporter. Export ins ANNIS- und PAULA-Format mit dem ANNISExporter und dem PAULAExporter (https://korpling.german.hu-berlin.de/saltnpepper/pepper/download/stable/Pepper_2019.06.11.zip)

Format:	PAULA-XML, ANNIS
Segmentierung der Annotation:	Eigenständig
Qualitätsprüfer:	Korpling
Datum:	2019
Qualitätsprüfung:	NA
Editor:	Korpling

4. Annotationsebenen – Annotationen zu linguistischen Eigenschaften

4.1. pos

Typ: *Layer* – pos

Name:	pos
Dokumentation:	http://www.ims.uni-stuttgart.de/projekte/corplex/TagSets/stts-table.html
Beschreibung:	Unter linguistischen Annotationen werden u. a. pos-Tagging, Lemmatisierung und Markierung von fremdsprachlichem Material zusammengefasst.

Typ: Preparationstep – pos

Schritt:	1
Aufbereitung:	Part-of-Speech-Tagging
Aufbereitungsart:	Automatisch
Tool:	Treetagger Version 3.2
Format:	Treetagger-Output
Segmentierung der Annotation:	basiert auf ‘norm’
Qualitätsprüfer:	Korpling
Datum:	2019
Qualitätsprüfung:	NA
Editor:	Korpling

Schritt:	2
Aufbereitung:	Überführung des TreeTagger-Outputs nach Excel 2013
Aufbereitungsart:	manuell
Tool:	Excel 2013
Format:	Excel 2013
Segmentierung der Annotation:	basiert auf ‘norm’
Qualitätsprüfer:	Korpling
Datum:	2019
Qualitätsprüfung:	NA
Editor:	Korpling

Schritt:	3
Aufbereitung:	Verbinden von leeren Zellen und der jeweils davor stehenden Zelle, die einen Wert enthält
Aufbereitungsart:	Automatisch
Tool:	SearchAndMerge.bas (https://hu.berlin/search-and-merge)
Format:	EXCEL
Segmentierung der Annotation:	basiert auf ‘norm’
Qualitätsprüfer:	Korpling
Datum:	2019
Qualitätsprüfung:	Korpling
Editor:	Catharina Fischer

Schritt:	4
Aufbereitung:	Konvertierung von Excel 2013 nach PAULA-XML und ANNIS
Aufbereitungsart:	Automatisch
Tool:	SaltNPepper, ExcelImporter (https://korpling.german.hu-berlin.de/saltnpepper/pepper/download/stable/Pepper_2019.06.11.zip)
Format:	PAULA-XML, ANNIS
Segmentierung der	

Annotation:	basiert auf ‘norm’
Qualitätsprüfer:	Korpling
Datum:	2019
Qualitätsprüfung:	Korpling
Editor:	Korpling

Typ: Annotationlayer – pos

Annotationsart:	Spannenannotation je norm-Einheit
Beschreibung:	Wortarten-Annotation nach dem Stuttgart-Tübingen-Tagset.

Typ: Annotationvalue – pos

Wert:	Wertbeschreibung:
STTS	Wortarten-Annotation nach dem Stuttgart-Tübingen-Tagset.

4.2. lemma

Typ: *Layer* – lemma

Name:	lemma
Dokumentation:	http://www.ims.uni-stuttgart.de/projekte/corplex/TreeTagger/
Beschreibung:	Unter linguistischen Annotationen werden u.a. pos-Tagging, Lemmatisierung und Markierung von fremdsprachlichem Material zusammengefasst.

Typ: Preparationstep – lemma

Schritt:	1
Aufbereitung:	Lemmatisierung
Aufbereitungsart:	Automatisch
Tool:	Treetagger Version 3.2

Format:	Treetagge-Output
Segmentierung der Annotation:	basiert auf ‘norm’
Qualitätsprüfer:	korpling
Datum:	2019
Qualitätsprüfung:	NA
Editor:	Carolin Odebrecht, Thomas Krause

Schritt:	2
Aufbereitung:	Konvertierung von Excel 2013 nach PAULA-XML und ANNIS
Aufbereitungsart:	automatisch
Tool:	Pepper, Konvertierung des Formats Excel nach ANNIS und PAULA mithilfe der folgenden Importer: SpreadsheetImporter. Export ins ANNIS- und PAULA-Format mit dem ANNISExporter und dem PAULAExporter (https://korpling.german.hu-berlin.de/saltnpepper/pepper/download/stable/Pepper_2019.06.11.zip)
Format:	PAULA-XML, ANNIS
Segmentierung der Annotation:	basiert auf ‘norm’
Qualitätsprüfer:	korpling
Datum:	2019
Qualitätsprüfung:	Korpling
Editor:	Korpling

Typ: Annotationlayer– lemma

Annotationsart:	Spannenannotation je norm-Einheit
Beschreibung:	Lemmatisierung: Normalisierung der flektierten Wortformen, Zuordnung der entsprechenden Lexikoneinträge zu jeder Wortform.

Typ: Annotationvalue – lemma

Wert:	Wertbeschreibung:
Normalisiertes Lemma	Lemmatisierung.

4.3. lang

Typ: *Layer* – lang

Name:	lang
Dokumentation:	Informationen zum gesamten Konvertierungsworkflow gibt es hier: https://github.com/langbank/conversion-pipeline
Beschreibung:	Unter linguistischen Annotationen werden u.a. pos-Tagging, Lemmatisierung und Markierung von fremdsprachlichem Material zusammengefasst.

Typ: Preparationstep – lang

Schritt:	1
Aufbereitung:	Linguistische Annotation
Aufbereitungsart:	manuell
Tool	Excel 2013
Format:	Excel 2013
Segmentierung der Annotation:	basiert auf 'dipl'
Qualitätsprüfer:	Catharina Fischer
Datum	2019
Qualitätsprüfung:	manuelle Korrektur
Editor:	Studierende der Humboldt-Universität zu Berlin und der Georg-August-Universität Göttingen

Schritt:	2
Aufbereitung:	Konvertierung von Excel 2013 nach PAULA-XML und ANNIS
Aufbereitungsart:	automatisch
Tool:	Pepper, Konvertierung des Formats Excel nach ANNIS und PAULA mithilfe der folgenden Importer: SpreadsheetImporter. Export ins ANNIS- und PAULA-Format mit dem ANNISExporter und dem PAULAExporter (https://korpling.german.hu-berlin.de/saltnpepper/pepper/download/stable/Pepper_2019.06.11.zip)
Format:	PAULA-XML, ANNIS

Segmentierung der Annotation:	basiert auf ‘dipl’
Qualitätsprüfer:	korpling
Datum:	2019
Qualitätsprüfung:	NA
Editor:	Korpling

Typ: Annotationlayer – lang

Annotationsart:	Annotationswert wird je dipl-Einhei zugewiesen
Beschreibung:	Kodiert die konkrete Sprache des Textes nach ISO Drei-Buchstaben-Code (z. B. lat).

Typ: Annotationvalue – lang

Wert:	Wertbeschreibung:
Kürzel der entsprechenden Spache	Getagged nach ISO Drei-Buchstaben-Code ISO 639-3 (http://www.sil.org/iso639-3/codes.asp?order=639_3&letter=%25)
deu	Deutsch (Default-Wert)
lat	Latein
grc	Altgriechisch
spa	Spanisch
ita	Italienisch
ara	Arabisch
fra	Französisch
pol	Polnisch
nld	Niederländisch/Belgisch/Flämisch
eng	Englisch
rus	Russisch
goh	Althochdeutsch
dan	Dänisch
hun	Ungarisch
egy	Ägyptisch

cze	Tschechisch/Böhmisch
swe	Schwedisch
nor	Norwegisch
undef	Für Abkürzungen, die nicht im Duden stehen
unknown	Für unknown (Unterstrich in ‘dipl’)

(Satz)zeichen und Ziffern bleiben immer <deu>.

Zweifelsfälle, bei denen nicht zwischen Schreibung oder Fremdsprache unterschieden werden kann, werden im Kontext beurteilt (Bspw. Capitel, Extract, October, Scribenten). „Deu“ wird annotiert, wenn das Wort in deutschem Kontext steht und „lat“, wenn das Wort in lateinischem Kontext steht. Namen bleiben „deu“ (Bspw. Brown, Plinius, Dioskurides, Loddike, Klusen, Lurk) - außer sie haben eine fremdsprachliche Flexionsendung. (Bspw. „lat“ für Plinii, Dioskuridi, Etmüllero)

Item/ Jtem sind „lat“

deu vs. lat: Wenn keine Entscheidung getroffen werden kann wird „lat“ verwendet.
lat vs. grc: Wenn keine Entscheidung getroffen werden kann, wird „lat“ verwendet.

Wenn etwas, laut Autor, in mehreren Sprachen den gleichen Namen hat, wir die erstgenannte Sprache verwendet.

4.4. abbr

Typ: Layer - abbr

Name:	abbr
Dokumentation:	Informationen zum gesamten Konvertierungsworkflow gibt es hier: https://github.com/langbank/conversion-pipeline.
Beschreibung:	Unter linguistischen Annotationen werden u.a. pos-Tagging, Lemmatisierung und Markierung von fremdsprachlichem Material zusammengefasst.

Typ: Preparationstep - abbr

Schritt:	1
Aufbereitung:	Linguistische Annotation
Aufbereitungsart:	manuell
Tool:	Excel 2013
Format:	Excel 2013
Segmentierung der Annotation:	basiert auf ‘norm’
Qualitätsprüfer:	Catharina Fischer

Datum:	2019
Qualitätsprüfung:	manuelle Korrektur
Editor:	Catharina Fischer

Schritt:	2
Aufbereitung:	Konvertierung von Excel 2013 nach PAULA-XML und ANNIS
Aufbereitungsart:	automatisch
Tool:	Pepper, Konvertierung des Formats Excel nach ANNIS und PAULA mithilfe der folgenden Importer: SpreadsheetImporter. Export ins ANNIS- und PAULA-Format mit dem ANNISExporter und dem PAULAExporter (https://korpling.german.hu-berlin.de/saltnpepper/pepper/download/stable/Pepper_2019.06.11.zip)
Format:	PAULA-XML, ANNIS
Segmentierung der Annotation:	basiert auf 'norm'
Qualitätsprüfer:	Catharina Fischer
Datum:	2019
Qualitätsprüfung:	NA
Editor:	Korpling

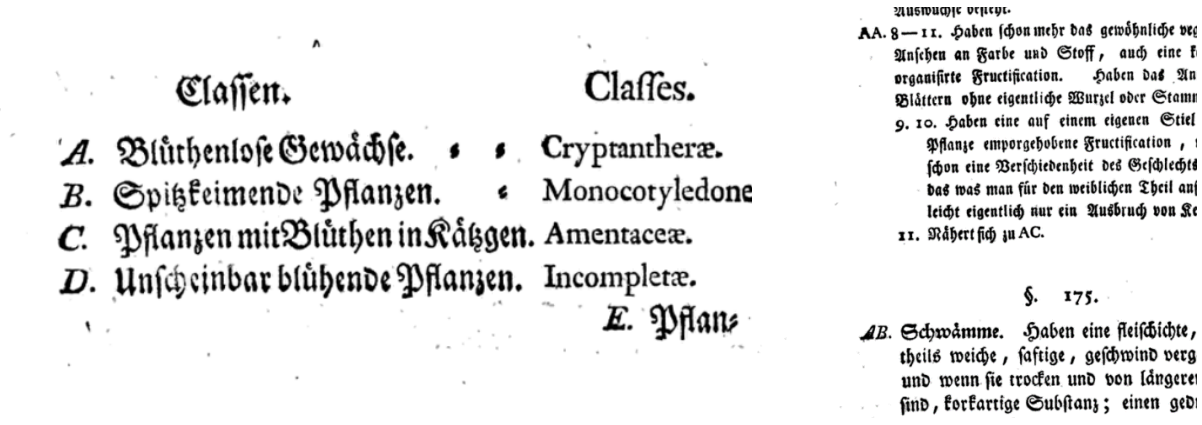
Typ: Annotationlayer – abbr

Annotationsart:	Annotationswert wird je norm-Einheit zugewiesen
Beschreibung:	Annotation von Ankürzungen in 'norm'

Typ: Annotationvalue – abbr

Wert:	Wertbeschreibung:
no	keine Abkürzung in 'norm' (Default-Wert)
yes	Abkürzung in 'norm'
unknown	für *unknown* in 'norm'

Aufzählungen, Mengenangaben (z.B. ij.) und Nummerierungen werden mit „no" annotiert (siehe Abb. unten). Im Zweifelsfall wird „yes" annotiert.



Classen. Classes.

A. Blüthenlose Gewächse. Cryptantheræ.
B. Spitzkeimende Pflanzen. Monocotyledone
C. Pflanzen mit Blüthen in Kätzgen. Amentaceæ.
D. Unscheinbar blühende Pflanzen. Incompletæ.
E. Pflan-

Auswuchse verlegt.
AA. 8—11. Haben schon mehr das gewöhnliche veg
Ansehen an Farbe und Stoff, auch eine f
organisirte Fructification. Haben das An
Blättern ohne eigentliche Wurzel oder Stamm
9. 10. Haben eine auf einem eigenen Stiel
Pflanze emporgehobene Fructification,
schon eine Verschiedenheit des Geschlechts
das was man für den weiblichen Theil an
leicht eigentlich nur ein Ausbruch von Ke
11. Nähert sich zu AC.

§. 175.

AB. Schwämme. Haben eine fleischichte,
theils weiche, saftige, geschwind verg
und wenn sie trocken und von längere
sind, korkartige Substanz; einen gedi

4.5. comp

Typ: *Layer* – komp

Name:	comp
Dokumentation:	Siehe Dokumentation von prot. Informationen zum gesamten Konvertierungsworkflow gibt es hier: https://github.com/langbank/conversion-pipeline.
Beschreibung:	Unter linguistischen Annotationen werden u.a. pos-Tagging, Lemmatisierung und Markierung von fremdsprachlichem Material zusammengefasst.

Typ: Preparationstep – comp

Schritt:	1
Aufbereitung:	Linguistische Annotation
Aufbereitungsart:	manuell
Tool	Excel 2010
Format:	Excel 2010
Segmentierung der Annotation:	basiert auf ‘dipl’
Qualitätsprüfer:	Laura Perlitz
Datum	2014
Qualitätsprüfung:	manuelle Korrektur
Editor:	Laura Perlitz

Schritt:	2
Aufbereitung:	Konvertierung von Excel 2010 nach Excel 2013
Aufbereitungsart:	manuell
Tool:	Excel 2013
Format:	Excel 2013
Segmentierung der Annotation:	basiert auf ‘dipl’
Qualitätsprüfer:	korpling
Datum:	2016

Qualitätsprüfung:	NA
Editor:	Carolin Odebrecht

Schritt:	3
Aufbereitung:	Konvertierung von Excel 2013 nach PAULA-XML und ANNIS
Aufbereitungsart:	automatisch
Tool:	Pepper, Konvertierung des Formats Excel nach ANNIS und PAULA mithilfe der folgenden Importer: SpreadsheetImporter. Export ins ANNIS- und PAULA-Format mit dem ANNISExporter und dem PAULAExporter (https://korpling.german.hu-berlin.de/saltnpepper/pepper/download/stable/Pepper_2019.06.11.zip)
Format:	PAULA-XML, ANNIS
Segmentierung der Annotation:	basiert auf 'dipl'
Qualitätsprüfer:	korpling
Datum:	2018
Qualitätsprüfung:	NA
Editor:	LAUDATIO, LangBank

Typ: Annotationlayer – comp

Annotationsart:	Spannenannotation über mehrere dipl-Einheiten
Beschreibung:	Markiert alle im Korpus enthaltenen zusammen, getrennt oder mit Bindestrich geschriebenen Komposita mit einem substantivischen Zweitglied.

Typ: Annotationvalue – comp

Wert:	Wertbeschreibung:
k	Kompositum mit substantivischem Zweitglied

4.6. comp_orth

Typ: *Layer* – comp_orth

Name:	comp_orth
Dokumentation:	Ist ein Kompositum an einer Morphemgrenze durch einen Zeilenumbruch unterbrochen, wurde es entweder mit „lb1“ oder „lb2“ markiert, je nachdem, ob ein Bindestrich vorhanden ist oder nicht. Ist dies nicht der Fall und das Kompositum wurde innerhalb einer Morphemgrenze unterbrochen, wird es als zusammen geschrieben markiert, da davon ausgegangen wird, dass der Autor das Kompositum nicht getrennt geschrieben hätte, wenn an dieser Stelle kein Zeilenumbruch vorkäme. Mit „bs“ wurden nur diejenigen Komposita markiert, die innerhalb einer Zeile mit einem Bindestrich versehen sind. Informationen zum gesamten Konvertierungsworkflow gibt es hier: https://github.com/langbank/conversion-pipeline.
Beschreibung:	Unter linguistischen Annotationen werden u.a. pos-Tagging, Lemmatisierung und Markierung von fremdsprachlichem Material zusammengefasst.

Typ: Preparationstep – comp_orth

Schritt:	1
Aufbereitung:	Linguistische Annotation
Aufbereitungsart:	manuell
Tool	Excel 2010
Format:	Excel 2010
Segmentierung der Annotation:	basiert auf ‘dipl’
Qualitätsprüfer:	Laura Perlitz
Datum	2014
Qualitätsprüfung:	manuelle Korrektur
Editor:	Laura Perlitz

Schritt:	2
Aufbereitung:	Konvertierung von Excel 2010 nach Excel 2013
Aufbereitungsart:	manuell
Tool:	Excel 2013
Format:	Excel 2013
Segmentierung der Annotation:	basiert auf ‘dipl’
Qualitätsprüfer:	korpling
Datum:	2016

Qualitätsprüfung:	NA
Editor:	Carolin Odebrecht
Schritt:	3
Aufbereitung:	Konvertierung von Excel 2013 nach PAULA-XML und ANNIS
Aufbereitungsart:	automatisch
Tool:	Pepper, Konvertierung des Formats Excel nach ANNIS und PAULA mithilfe der folgenden Importer: SpreadsheetImporter. Export ins ANNIS- und PAULA-Format mit dem ANNISExporter und dem PAULAExporter https://korpling.german.hu-berlin.de/saltnpepper/pepper/download/stable/Pepper_2019.06.11.zip)
Format:	PAULA-XML, ANNIS
Segmentierung der Annotation:	basiert auf 'dipl'
Qualitätsprüfer:	korpling
Datum:	2018
Qualitätsprüfung:	NA
Editor:	LAUDATIO, LangBank

Typ: Annotationlayer – comp_orth

Annotationsart:	Spannenannotation über mehrere dipl-Einheiten
Beschreibung:	Beschreibt die Orthographie der in der Ebene „komp" annotierten Komposita.

Typ: Annotationvalue – comp_orth

Wert:	Wertbeschreibung:
zs	zusammen geschrieben
gtr	getrennt geschrieben
bs	Bindestrichmarkierung (innerhalb einer Zeile)
lb1	unterbrochen durch Zeilenumbruch (ohne Bindestrichmarkierung)
lb2	unterbrochen durch Zeilenumbruch (mit Bindestrichmarkierung)

4.7. prot

Typ: *Layer* – prot

Name:	prot
Dokumentation:	Der Prototyp 1 (prot1) markiert alle Token, die eindeutig als Komposita mit substantivischem Zweitglied zu identifizieren sind. Folgende Kriterien müssen dabei erfüllt sein: Es liegen mindestens zwei freie Morpheme vor, die entweder zusammen oder mit Bindestrich geschrieben wurden. Es darf wortintern keine Flexion auftreten und das Kompositum kann syntaktisch nicht erweitert werden, beispielsweise durch Adjektiv- oder Präpositionalphrasen (z.B. *ſtabwurz*). Der Prototyp 2 (prot2) fasst alle Tokens zusammen, die sehr wahrscheinlich Komposita sind. Diese bestehen ebenfalls aus zwei freien Morphemen, werden aber getrennt geschrieben oder wurden durch einen Zeilenumbruch unterbrochen und enthalten keine Bindestrichmarkierung. Sie dürfen wie Komposita des Prototyp 1 wortintern nicht flektierbar und syntaktisch nicht erweiterbar sein (z.B. *wermut ſaft*). Der Prototyp 3 (prot3) markiert alle Zweifelsfälle, also Fälle, die weder eindeutig als Phrase noch als Kompositum zu identifizieren sind. Wieder liegen zwei freie Morpheme sowie Getrenntschreibung vor, jedoch kann man keine genaue Aussage darüber treffen, ob eine wortinterne Flexion vorgenommen wurde oder möglich wäre und ob das jeweilige Vorkommen syntaktisch erweiterbar ist (z.B. *widers Teuffels Biſſz*). Die Zweifelsfälle wurden nicht in der Ebene “komp” aufgenommen. Informationen zum gesamten Konvertierungsworkflow gibt es hier: https://github.com/langbank/conversion-pipeline.
Beschreibung:	Unter linguistischen Annotationen werden u.a. pos-Tagging, Lemmatisierung und Markierung von fremdsprachlichem Material zusammengefasst.

Typ: Preparationstep – prot

Schritt:	1
Aufbereitung:	Linguistische Annotation
Aufbereitungsart:	manuell
Tool	Excel 2010
Format:	Excel 2010
Segmentierung der	basiert auf ‘dipl’

Annotation:	
Qualitätsprüfer:	Laura Perlitz
Datum	2014
Qualitätsprüfung:	manuelle Korrektur
Editor:	Laura Perlitz

Schritt:	2
Aufbereitung:	Konvertierung von Excel 2010 nach Excel 2013
Aufbereitungsart:	manuell
Tool:	Excel 2013
Format:	Excel 2013
Segmentierung der Annotation:	basiert auf 'dipl'
Qualitätsprüfer:	korpling
Datum:	2016
Qualitätsprüfung:	NA
Editor:	Carolin Odebrecht

Schritt:	3
Aufbereitung:	Konvertierung von Excel 2013 nach PAULA-XML und ANNIS
Aufbereitungsart:	automatisch
Tool:	Pepper, Konvertierung des Formats Excel nach ANNIS und PAULA mithilfe der folgenden Importer: SpreadsheetImporter. Export ins ANNIS- und PAULA-Format mit dem ANNISExporter und dem PAULAExporter (https://korpling.german.hu-berlin.de/saltnpepper/pepper/download/stable/Pepper_2019.06.11.zip)
Format:	PAULA-XML, ANNIS
Segmentierung der Annotation:	basiert auf 'dipl'
Qualitätsprüfer:	korpling
Datum:	2018
Qualitätsprüfung:	NA
Editor:	LAUDATIO, LangBank

Typ: Annotationlayer – prot

Annotationsart:	Spannenannotation über mehrere dipl-Einheiten

Beschreibung:	Weist jedem Wert in der Ebene Kompositum einen Prototypen zu. Der Prototyp gibt an, wie wahrscheinlich ein Vorkommen ein Substantivkompositum ist.

Typ: Annotationvalue – prot

Wert:	Wertbeschreibung:
prot1	sicher als Kompositum zu identifizieren
prot2	sehr wahrscheinlich Kompositum
prot3	Zweifelsfall; nicht in der Ebene „komp“ markiert

4.8. comp_n

Typ: *Layer* – comp_n

Name:	comp_n
Dokumentation:	Morphologische Annotation von Komposita mit nominalem Zweitglied. Hierin können auch getrenntgeschriebene Komposita vorkommen, die unter bestimmten Umständen auch als Konstruktionen mit pränominalem Genitiv interpretiert werden können. Da jedoch konkrete Anhaltspunkte vorliegen, diese als Komposita zu interpretieren (z.B. Modifikatoren, Flexions-/ Fugenelemente), wurden sie in dieser Ebene annotiert. Fälle, die weder als Kompositium noch als Syntagma eingeordnet werden konnten, wurden in comp_amb annotiert. Informationen zum gesamten Konvertierungsworkflow gibt es hier: https://github.com/langbank/conversion-pipeline.
Beschreibung:	Unter linguistischen Annotationen werden u.a. pos-Tagging, Lemmatisierung und Markierung von fremdsprachlichem Material zusammengefasst.

Typ: Preparationstep – comp_n

Schritt:	1
Aufbereitung:	Linguistische Annotation
Aufbereitungsart:	manuell
Tool	Excel 2013
Format:	Excel 2013

Segmentierung der Annotation:	basiert auf ‘dipl’
Qualitätsprüfer:	Laura Perlitz
Datum	2018
Qualitätsprüfung:	manuelle Korrektur
Editor:	Laura Perlitz

Schritt:	2
Aufbereitung:	Konvertierung von Excel 2013 nach PAULA-XML und ANNIS
Aufbereitungsart:	automatisch
Tool:	Pepper, Konvertierung des Formats Excel nach ANNIS und PAULA mithilfe der folgenden Importer: SpreadsheetImporter. Export ins ANNIS- und PAULA-Format mit dem ANNISExporter und dem PAULAExporter (https://korpling.german.hu-berlin.de/saltnpepper/pepper/download/stable/Pepper_2019.06.11.zip)
Format:	PAULA-XML, ANNIS
Segmentierung der Annotation:	basiert auf ‘dipl’
Qualitätsprüfer:	korpling
Datum:	2018
Qualitätsprüfung:	NA
Editor:	LAUDATIO, LangBank

Typ: Annotationlayer – comp_n

Annotationsart:	Spannenannotation über mehrere dipl-Einheiten
Beschreibung:	Markiert alle im Korpus enthaltenen zusammen, getrennt oder mit Bindestrich geschriebenen Komposita mit einem substantivischen Zweitglied.

Typ: Annotationvalue – comp_n

Wert:	Wertbeschreibung:
	Additive Werte zur Beschreibung der morphologischen Struktur der nominalen Komposita, z.B. N_N für Komposita, die aus zwei Nomina bestehen. Suffixe wurden

	nur berücksichtigt, wenn Suffigierung der letzte Wortbildungsprozess war ([[*Kindbett*]*erin*]). Zwischen den Kürzeln für die Kategorien der einzelnen morphologischen Elemente stehen Unterstriche. Ist die Kategorie eines Elements nicht ermittelbar, wird der Platzhalter „X“ verwendet.
N	Nomen, z.B. *Ampfer* in *Ampffer kraut*
A	Adjektiv, z.B. *weiß* in *Weiſs kol*
V	Verb, z.B. *distillieren* in *diſti⸗ lierGlåsern*
ADV	Adverb, z.B. *aber* in *Aberglauben*
APPR	Präposition, z.B. *bei* in *beynammen*
CARD	Kardinalzahl, z.B. drei in treyfuſz
SUFF	Suffixe (bei mehreren hintereinnder wird nur eins gezählt, z.B. *er-in* in *kindbetterin*
CONV	Konversion, z.B. bei *das hertz zittern*
X	Platzhalter für intransparente Wortteile, z.B. *schel* in *Schelwurtz*

4.9. comp_n_graph

Typ: *Layer* – comp_n_graph

Name:	comp_n_graph
Dokumentation:	Beschreibung der Graphematik der Komposita mit nominalem Zweitglied. Ist ein Kompositum an einer Morphemgrenze durch einen Zeilenumbruch unterbrochen, wurde es entweder mit „lb1“ oder „lb2“ markiert, je nachdem, ob ein Bindestrich vorhanden ist oder nicht. Ist dies nicht der Fall und das Kompositum wurde innerhalb einer Morphemgrenze unterbrochen, wird es als zusammen geschrieben markiert, da davon ausgegangen wird, dass der Autor das Kompositum nicht getrennt geschrieben hätte, wenn an dieser Stelle kein Zeilenumbruch vorkäme. Informationen zum gesamten Konvertierungsworkflow gibt es hier: https://github.com/langbank/conversion-pipeline.
Beschreibung:	Unter linguistischen Annotationen werden u.a. pos-Tagging, Lemmatisierung und Markierung von fremdsprachlichem Material zusammengefasst.

Typ: Preparationstep – comp_n_graph

Schritt:	1
Aufbereitung:	Linguistische Annotation
Aufbereitungsart:	manuell
Tool	Excel 2013
Format:	Excel 2013
Segmentierung der Annotation:	basiert auf 'dipl'
Qualitätsprüfer:	Laura Perlitz
Datum	2018
Qualitätsprüfung:	manuelle Korrektur
Editor:	Laura Perlitz

Schritt:	2
Aufbereitung:	Konvertierung von Excel 2013 nach PAULA-XML und ANNIS
Aufbereitungsart:	automatisch
Tool:	Pepper, Konvertierung des Formats Excel nach ANNIS und PAULA mithilfe der folgenden Importer: SpreadsheetImporter. Export ins ANNIS- und PAULA-Format mit dem ANNISExporter und dem PAULAExporter (https://korpling.german.hu-berlin.de/saltnpepper/pepper/download/stable/Pepper_2019.06.11.zip)
Format:	PAULA-XML, ANNIS
Segmentierung der Annotation:	basiert auf 'dipl'
Qualitätsprüfer:	korpling
Datum:	2018
Qualitätsprüfung:	NA
Editor:	LAUDATIO, LangBank

Typ: Annotationlayer – comp_n_graph

Annotationsart:	Spannenannotation über mehrere dipl-Einheiten

Beschreibung:	Beschreibt die Graphematik der in der Ebene „comp_n“ annotierten Komposita.

Typ: Annotationvalue – comp_n_graph

Wert:	Wertbeschreibung:
nospace	zusammen geschrieben
sep	getrennt geschrieben
hyph	Bindestrichmarkierung (innerhalb einer Zeile)
lb1	unterbrochen durch Zeilenumbruch (ohne Bindestrichmarkierung)
lb2	unterbrochen durch Zeilenumbruch (mit Bindestrichmarkierung)
camel	Binnengroßschreibung

Komposita, die aus mehr als zwei lexikalischen Teilen entstehen, erhalten einen additiven Wert (verbunden durch Unterstriche), z.B. „nospace_sep“ bei *Saurampffer waffer*.

4.10. comp_n_mod

Typ: *Layer* – comp_n_mod

Name:	comp_n_mod
Dokumentation:	Gibt an, ob die in der Ebene <comp_n> annotierten potentiellen Komposita einen oder mehrere(flektierende) Modifikatoren erhalten haben und welche(n).
Beschreibung:	Unter linguistischen Annotationen werden u.a. pos-Tagging, Lemmatisierung und Markierung von fremdsprachlichem Material zusammengefasst.

Typ: Preparationstep – comp_n_mod

Schritt:	1
Aufbereitung:	Linguistische Annotation
Aufbereitungsart:	manuell

Tool	Excel 2013
Format:	Excel 2013
Segmentierung der Annotation:	basiert auf ‘dipl’
Qualitätsprüfer:	Laura Perlitz
Datum	2018
Qualitätsprüfung:	manuelle Korrektur
Editor:	Laura Perlitz

Schritt:	2
Aufbereitung:	Konvertierung von Excel 2013 nach PAULA-XML und ANNIS
Aufbereitungsart:	automatisch
Tool:	Pepper, Konvertierung des Formats Excel nach ANNIS und PAULA mithilfe der folgenden Importer: SpreadsheetImporter. Export ins ANNIS- und PAULA-Format mit dem ANNISExporter und dem PAULAExporter https://korpling.german.hu-berlin.de/saltnpepper/pepper/download/stable/Pepper_2019.06.11.zip)
Format:	PAULA-XML, ANNIS
Segmentierung der Annotation:	basiert auf ‘dipl’
Qualitätsprüfer:	korpling
Datum:	2018
Qualitätsprüfung:	NA
Editor:	LAUDATIO, LangBank

Typ: Annotationlayer – comp_n_mod

Annotationsart:	Spannenannotation über mehrere dipl-Einheiten
Beschreibung:	Gibt an, ob die in der Ebene <comp_n> annotierten potentiellen Komposita einen oder mehrere (flektierende) Modifikatoren erhalten haben und welche(n).

Typ: Annotationvalue – comp_n_mod

Wert:	Wertbeschreibung:
	Dem jeweiligen (potentiellen) Kompositum wurde ein Wert aus dem STTS (in Kleinbuchstaben) zugewiesen, der die Wortart des Modifikators beschreibt. Bei mehreren Modifikatoren wurden additive Werte vergeben, die durch einen Unterstrich separiert sind, z.B. 'art_adja'.
art	bestimmter oder unbestimmter Artikel
apprart	Präposition mit Artikel
adja	attributives Adjektiv
piat	attribuierendes Indefinitpronomen ohne Determiner (Determinierer wurden separat durch "art" oder "apprart" erfasst)
pposat	attribuierendes Possessivpronomen
pdat	attribuierendes Demonstrativpronomen
prelat	attribuierendes Relativpronomen
card	Kardinalzahl
NA	keine Angabe möglich
n	kein Modifikator vorhanden

4.11. comp_a

Typ: *Layer* – comp_a

Name:	comp_a
Dokumentation:	Beschreibung der Graphematik der Komposita mit nominalem Zweitglied. Ist ein Kompositum an einer Morphemgrenze durch einen Zeilenumbruch unterbrochen, wurde es entweder mit „lb1“ oder „lb2“ markiert, je nachdem, ob ein Bindestrich vorhanden ist oder nicht. Ist dies nicht der Fall und das Kompositum wurde innerhalb einer Morphemgrenze unterbrochen, wird es als zusammen geschrieben markiert, da davon ausgegangen wird, dass der Autor das Kompositum nicht getrennt geschrieben hätte, wenn an dieser Stelle kein Zeilenumbruch vorkäme. Informationen zum gesamten Konvertierungsworkflow gibt es hier: https://github.com/langbank/conversion-pipeline.
Beschreibung:	Unter linguistischen Annotationen werden u.a. pos-Tagging, Lemmatisierung und Markierung von fremdsprachlichem Material zusammengefasst.

Typ: Preparationstep – comp_a

Schritt:	1
Aufbereitung:	Linguistische Annotation
Aufbereitungsart:	manuell
Tool	Excel 2013
Format:	Excel 2013
Segmentierung der Annotation:	basiert auf 'dipl'
Qualitätsprüfer:	Laura Perlitz
Datum	2018
Qualitätsprüfung:	manuelle Korrektur
Editor:	Laura Perlitz

Schritt:	2
Aufbereitung:	Konvertierung von Excel 2013 nach PAULA-XML und ANNIS
Aufbereitungsart:	automatisch
Tool:	Pepper, Konvertierung des Formats Excel nach ANNIS und PAULA mithilfe der folgenden Importer: SpreadsheetImporter. Export ins ANNIS- und PAULA-Format mit dem ANNISExporter und dem PAULAExporter (https://korpling.german.hu-berlin.de/saltnpepper/pepper/download/stable/Pepper_2019.06.11.zip)
Format:	PAULA-XML, ANNIS
Segmentierung der Annotation:	basiert auf 'dipl'
Qualitätsprüfer:	korpling
Datum:	2018
Qualitätsprüfung:	NA
Editor:	LAUDATIO, LangBank

Typ: Annotationlayer – comp_a

Annotationsart:	Spannenannotation über mehrere dipl-Einheiten
Beschreibung:	Beschreibt die Orthographie der in der Ebene „komp“ annotierten Komposita.

Typ: Annotationvalue – comp_a

Wert:	Wertbeschreibung:
	Additive Werte zur Beschreibung der morphologischen Struktur der nominalen Komposita, z.B. A_A für Komposita, die aus zwei Adjektiven bestehen. Suffixe wurden nur berücksichtigt, wenn Suffigierung der letzte Wortbildungsprozess war. Zwischen den Kürzeln für die Kategorien der einzelnen morphologischen Elemente stehen Unterstriche.
N	Nomen, z.B. *Asche* in *aſchenfarben*
A	Adjektiv, z.B. *wohl* in *wohlriechend*
ADV	Adverb, z.B. *aber* in *Aberglauben*
CARD	Kardinalzahl, z.B. *vier* in *viereckicht*
SUFF	Suffixe (bei mehreren hintereinnder wird nur eins gezählt, z.B. *-isch* in *Waldenburgiſchen*
CONV	Konversion, z.B. bei *die Hochgelehrten*
farb	Das letzte Glied enthält die Wurzel „farb“, z.B. *braunfarb*.

4.12. comp_a_graph

Typ: *Layer* – comp_a_graph

Name:	comp_a_graph
Dokumentation:	Beschreibung der Graphematik der Komposita mit nominalem Zweitglied. Ist ein Kompositum an einer Morphemgrenze durch einen Zeilenumbruch unterbrochen, wurde es entweder mit „lb1“ oder „lb2“ markiert, je nachdem, ob ein Bindestrich vorhanden ist oder nicht. Ist dies nicht der Fall und das Kompositum wurde innerhalb einer Morphemgrenze unterbrochen, wird es als zusammen geschrieben markiert, da davon ausgegangen wird, dass der Autor das Kompositum nicht getrennt geschrieben hätte, wenn an dieser Stelle kein Zeilenumbruch vorkäme. Informationen zum gesamten Konvertierungsworkflow gibt es hier: https://github.com/langbank/conversion-pipeline.

Beschreibung:	Unter linguistischen Annotationen werden u.a. pos-Tagging, Lemmatisierung und Markierung von fremdsprachlichem Material zusammengefasst.

Typ: Preparationstep – comp_a_graph

Schritt:	1
Aufbereitung:	Linguistische Annotation
Aufbereitungsart:	manuell
Tool	Excel 2013
Format:	Excel 2013
Segmentierung der Annotation:	basiert auf 'dipl'
Qualitätsprüfer:	Laura Perlitz
Datum	2018
Qualitätsprüfung:	manuelle Korrektur
Editor:	Laura Perlitz

Schritt:	2
Aufbereitung:	Konvertierung von Excel 2013 nach PAULA-XML und ANNIS
Aufbereitungsart:	automatisch
Tool:	Pepper, Konvertierung des Formats Excel nach ANNIS und PAULA mithilfe der folgenden Importer: SpreadsheetImporter. Export ins ANNIS- und PAULA-Format mit dem ANNISExporter und dem PAULAExporter (https://korpling.german.hu-berlin.de/saltnpepper/pepper/download/stable/Pepper_2019.06.11.zip)
Format:	PAULA-XML, ANNIS
Segmentierung der Annotation:	basiert auf 'dipl'
Qualitätsprüfer:	korpling
Datum:	2018
Qualitätsprüfung:	NA
Editor:	LAUDATIO, LangBank

Typ: Annotationlayer – comp_a_graph

Annotationsart:	Spannenannotation über mehrere dipl-Einheiten
Beschreibung:	Beschreibt die Orthographie der in der Ebene „comp_a“ annotierten Komposita.

Typ: Annotationvalue – comp_a_graph

Wert:	Wertbeschreibung:
nospace	zusammen geschrieben
sep	getrennt geschrieben
hyph	Bindestrichmarkierung (innerhalb einer Zeile)
lb1	unterbrochen durch Zeilenumbruch (ohne Bindestrichmarkierung)
lb2	unterbrochen durch Zeilenumbruch (mit Bindestrichmarkierung)

4.13. comp_amb

Typ: *Layer* – comp_amb

Name:	comp_amb
Dokumentation:	Annotation von Wortfolgen, die sowohl als Komposita als auch als Syntagmen interpretiert werden können und bei denen der Kontext oder die Morphologie keine Hinweise für die Einordnung in die eine oder andere Kategorie liefern. Informationen zum gesamten Konvertierungsworkflow gibt es hier: https://github.com/langbank/conversion-pipeline.
Beschreibung:	Unter linguistischen Annotationen werden u.a. pos-Tagging, Lemmatisierung und Markierung von fremdsprachlichem Material zusammengefasst.

Typ: Preparationstep – comp_amb

Schritt:	1
Aufbereitung:	Linguistische Annotation
Aufbereitungsart:	manuell
Tool	Excel 2013
Format:	Excel 2013
Segmentierung der Annotation:	basiert auf ‘dipl’
Qualitätsprüfer:	Laura Perlitz
Datum	2018
Qualitätsprüfung:	manuelle Korrektur
Editor:	Laura Perlitz

Schritt:	2
Aufbereitung:	Konvertierung von Excel 2013 nach PAULA-XML und ANNIS
Aufbereitungsart:	automatisch
Tool:	Pepper, Konvertierung des Formats Excel nach ANNIS und PAULA mithilfe der folgenden Importer: SpreadsheetImporter. Export ins ANNIS- und PAULA-Format mit dem ANNISExporter und dem PAULAExporter (https://korpling.german.hu-berlin.de/saltnpepper/pepper/download/stable/Pepper_2019.06.11.zip)
Format:	PAULA-XML, ANNIS
Segmentierung der Annotation:	basiert auf ‘dipl’
Qualitätsprüfer:	korpling
Datum:	2018
Qualitätsprüfung:	NA
Editor:	LAUDATIO, LangBank

Typ: Annotationlayer – comp_amb

Annotationsart:	Spannenannotation über mehrere dipl-Einheiten

Beschreibung:	Annotation von Wortfolgen, die möglicherweise Komposita sein könnten.

Typ: Annotationvalue – comp_amb

Wert:	Wertbeschreibung:
a	Adjektiv-Nomen-Abfolgen, bei denen das Adjektiv unflektiert ist und bei denen man im modernen Deutschen eher ein Kompositum verwenden würde (z.B. das edel geſteine). Es gibt noch weitere AN-Abfolgen, die ebenfalls ein unflektiertes Adjektiv enthalten, die aber im heutigen Deutschen auch genauso (oder noch eher) als attributive Adjektive verwendet werden können. Diese wurden jedoch in ADJA_uninfl annotiert (z.B.).
gpre	Nomen-Nomen-Abfolgen, die sowohl als Komposita als auch als Nomina mit pränominalem Genitivattribut interpretiert werden können und bei denen der Kontext oder die Flexion keinen Aufschluss auf den Status geben können.

4.14. infl_fuge

Typ: *Layer* – infl_fuge

Name:	infl_fuge
Dokumentation:	Annotation der (Nicht-) Existenz von Fugenelementen, bzw. Flexionselementen zwischen zwei oder mehr lexikalischen Einheiten, die potentielle Komposita darstellen (v.a. N-N-Abfolgen). Informationen zum gesamten Konvertierungsworkflow gibt es hier: https://github.com/langbank/conversion-pipeline.
Beschreibung:	Unter linguistischen Annotationen werden u.a. pos-Tagging, Lemmatisierung und Markierung von fremdsprachlichem Material zusammengefasst.

Typ: Preparationstep – infl_fuge

Schritt:	1
Aufbereitung:	Linguistische Annotation

Aufbereitungsart:	manuell
Tool	Excel 2013
Format:	Excel 2013
Segmentierung der Annotation:	basiert auf ‘dipl’
Qualitätsprüfer:	Laura Perlitz
Datum	2018
Qualitätsprüfung:	manuelle Korrektur
Editor:	Laura Perlitz

Schritt:	2
Aufbereitung:	Konvertierung von Excel 2013 nach PAULA-XML und ANNIS
Aufbereitungsart:	automatisch
Tool:	Pepper, Konvertierung des Formats Excel nach ANNIS und PAULA mithilfe der folgenden Importer: SpreadsheetImporter. Export ins ANNIS- und PAULA-Format mit dem ANNISExporter und dem PAULAExporter (https://korpling.german.hu-berlin.de/saltnpepper/pepper/download/stable/Pepper_2019.06.11.zip)
Format:	PAULA-XML, ANNIS
Segmentierung der Annotation:	basiert auf ‘dipl’
Qualitätsprüfer:	korpling
Datum:	2018
Qualitätsprüfung:	NA
Editor:	LAUDATIO, LangBank

Typ: Annotationlayer – infl_fuge

Annotationsart:	Spannenannotation über mehrere dipl-Einheiten
Beschreibung:	Annotation von Flexion bzw. Fugenelementen bei Wortfolgen, die möglicherweise Komposita sein könnten.

Typ: Annotationvalue – infl_fuge

Wert:	Wertbeschreibung:
y	Es ist ein Fugenelement, bzw. ein Flexionselement vorhanden (‚yes').
n	Es ist kein Fugenelement, bzw. Flexionselement vorhanden (‚no').
yn	Gilt für potentielle Komposita mit mehr als zwei lexikalischen Wortteilen, die sowohl mit als auch ohne Fugenelement oder Flexionselement realisiert wurden (z.B. *Jungkfraw**en** har*)
NA	Es kann nicht beurteilt werden, ob ein Fugenelement, bzw. Flexionselement vorhanden ist (meistens weil das betroffene Wort/lexikalische Material nicht mehr transparent ist oder aus einer anderen Sprache (Latein) stammt).

4.15. comp_lex

Typ: *Layer* – comp_lex

Name:	comp_lex
Dokumentation:	Kategorisierung der (potentiellen) Komposita in lexikalisierte und nicht-lexikalisierte Komposita nach bestimmten Kriterien (s. unten). Informationen zum gesamten Konvertierungsworkflow gibt es hier: https://github.com/langbank/conversion-pipeline.
Beschreibung:	Unter linguistischen Annotationen werden u.a. pos-Tagging, Lemmatisierung und Markierung von fremdsprachlichem Material zusammengefasst.

Typ: Preparationstep – comp_lex

Schritt:	1
Aufbereitung:	Linguistische Annotation
Aufbereitungsart:	manuell
Tool	Excel 2013

Format:	Excel 2013
Segmentierung der Annotation:	basiert auf ‘dipl’
Qualitätsprüfer:	Laura Perlitz
Datum	2018
Qualitätsprüfung:	manuelle Korrektur
Editor:	Laura Perlitz

Schritt:	2
Aufbereitung:	Konvertierung von Excel 2013 nach PAULA-XML und ANNIS
Aufbereitungsart:	automatisch
Tool:	Pepper, Konvertierung des Formats Excel nach ANNIS und PAULA mithilfe der folgenden Importer: SpreadsheetImporter. Export ins ANNIS- und PAULA-Format mit dem ANNISExporter und dem PAULAExporter (https://korpling.german.hu-berlin.de/saltnpepper/pepper/download/stable/Pepper_2019.06.11.zip)
Format:	PAULA-XML, ANNIS
Segmentierung der Annotation:	basiert auf ‘dipl’
Qualitätsprüfer:	korpling
Datum:	2018
Qualitätsprüfung:	NA
Editor:	LAUDATIO, LangBank

Typ: Annotationlayer – comp_lex

Annotationsart:	Spannenannotation über mehrere dipl-Einheiten
Beschreibung:	Annotation der Lexikalisierung von Wortfolgen, die möglicherweise Komposita sein könnten.

Typ: Annotationvalue – comp_lex

Wert:	Wertbeschreibung:

lex	lexikalisierte Komposita, die nicht (mehr) als Syntagma ausgedrückt werden können, weil ihre addierten Teilbedeutungen stark von der Kompositionsbedeutung abweichen (betroffene Themenbereiche: Pflanzennamen, geographische Namen, bestimmte Krankheiten, Körperteile, Tierarten, Sternzeichen), aber nicht wenn sie mit einem zusätzlichen Glied versehen sind, das nicht zu dem lexikalisierten Wort gehört, z.B. *Eisenkrautsaft oder Beifußblumen* (vs. Johannisblumen), *Blutwassersucht.* Verbindungen mit "Baum" werden wie Verbindungen mit "Kraut" behandelt - meistens sind sie ein zusätzliches Glied zur eigentlichen Pflanzenbezeichnung - *Kirsche -> Kirschbaum*, Eiche -> *Eichenbaum*; manchmal gehören sie aber zur lexikalisierten Pflanzenbezeichnung (z.B. *Schildkraut, Rutelkraut, Wunderbaum*)
n	nicht-lexikalisierte (potentielle) Komposita, die nicht den oben genannten Kriterien entsprechen

Einzelentscheidungen

Vorkommen	**lex**	**n**
Jungfrau	x	
Liebhaber	x	
Wohlfahrt	x	
Mehrteil		x
Vaterland	x	
Bauchfluss		x
Aberglaube	x	
Aderlass	x	
Brachmonat	x	
Brosam	x	
Erdapfel	x	
Gutdünken	x	
Maibutter		x
Notdurft	x	
Freiherr	x	
Gegenteil	x	
Nachteil	x	
Ohnmacht	x	
Gliedmaßen	x	
Nachgeburt	x	
Widerwille	x	

4.16. adja_uninfl

Typ: *Layer* – adja_uninfl

Name:	adja_uninfl
Dokumentation:	Annotation von unflektierten Adjektiven, die direkt vor einem Nomen stehen. Stehen vor einem Nomen mehrere unflektierte Adjektive, wird nur dasjenige annotiert, das unmittelbar vor dem Nomen steht. Informationen zum gesamten Konvertierungsworkflow gibt es hier: https://github.com/langbank/conversion-pipeline.
Beschreibung:	Unter linguistischen Annotationen werden u.a. pos-Tagging, Lemmatisierung und Markierung von fremdsprachlichem Material zusammengefasst.

Typ: Preparationstep – adja_uninfl

Schritt:	1
Aufbereitung:	Linguistische Annotation
Aufbereitungsart:	manuell
Tool	Excel 2013
Format:	Excel 2013
Segmentierung der Annotation:	basiert auf 'norm'
Qualitätsprüfer:	Laura Perlitz
Datum	2018
Qualitätsprüfung:	manuelle Korrektur
Editor:	Laura Perlitz

Schritt:	2
Aufbereitung:	Konvertierung von Excel 2013 nach PAULA-XML und ANNIS
Aufbereitungsart:	automatisch
Tool:	Pepper, Konvertierung des Formats Excel nach ANNIS und PAULA mithilfe der folgenden Importer: SpreadsheetImporter. Export ins ANNIS- und PAULA-Format mit dem ANNISExporter und dem PAULAExporter (https://korpling.german.hu-berlin.de/saltnpepper/pepper/download/stable/Pepper_2019.06.11.zip)
Format:	PAULA-XML, ANNIS

Segmentierung der Annotation:	basiert auf ‘norm’
Qualitätsprüfer:	korpling
Datum:	2018
Qualitätsprüfung:	NA
Editor:	LAUDATIO, LangBank

Typ: Annotationlayer – adja_uninfl

Annotationsart:	Spannenannotation für einzelne norm-Einheiten
Beschreibung:	Annotation von Wortfolgen, die möglicherweise Komposita sein könnten.

Typ: Annotationvalue – adja_uninfl

Wert:	Wertbeschreibung:
uninfl	unflektierte Adjektive, die direkt vor einem Nomen stehen; stehen vor einem Nomen mehrere unflektierte Adjektive, wird nur dasjenige annotiert, das unmittelbar vor dem Nomen steht

4.17. attr_gen

Typ: *Layer* – attr_gen

Name:	attr_gen
Dokumentation:	Das Bezugsnomen und das Genitivattribut wurden in ein und derselben Spanne annotiert. Das Gleiche gilt für verschachtelte sowie koordinierte und durch weitere Phrasen erweiterte Genitivkonstruktionen (z.B. *Beobachtungen der nach der Bestäubung in den vegetabilischen Eyern erfolgenden Veränderungen*). Informationen zum gesamten Konvertierungsworkflow gibt es hier: https://github.com/langbank/conversion-pipeline.
Beschreibung:	Unter linguistischen Annotationen werden u.a. pos-Tagging, Lemmatisierung und Markierung von fremdsprachlichem Material zusammengefasst.

Typ: Preparationstep – attr_gen

Schritt:	1
Aufbereitung:	Linguistische Annotation
Aufbereitungsart:	manuell
Tool	Excel 2010
Format:	Excel 2010
Segmentierung der Annotation:	basiert auf 'dipl'
Qualitätsprüfer:	Laura Perlitz
Datum	2018
Qualitätsprüfung:	manuelle Korrektur
Editor:	Laura Perlitz

Schritt:	2
Aufbereitung:	Konvertierung von Excel 2013 nach PAULA-XML und ANNIS
Aufbereitungsart:	automatisch
Tool:	Pepper, Konvertierung des Formats Excel nach ANNIS und PAULA mithilfe der folgenden Importer: SpreadsheetImporter. Export ins ANNIS- und PAULA-Format mit dem ANNISExporter und dem PAULAExporter (https://korpling.german.hu-berlin.de/saltnpepper/pepper/download/stable/Pepper_2019.06.11.zip)
Format:	PAULA-XML, ANNIS
Segmentierung der Annotation:	basiert auf 'dipl'
Qualitätsprüfer:	korpling
Datum:	2018
Qualitätsprüfung:	NA
Editor:	LAUDATIO, LangBank

Typ: Annotationlayer – attr_gen

Annotationsart:	Spannenannotation über mehrere dipl-Einheiten

Beschreibung:	Markiert alle im Korpus enthaltenen Nominalphrasen, die ein Genitivattribut enthalten.

Typ: Annotationvalue – attr_gen

Wert:	Wertbeschreibung:
gpre	pränominales Genitivattribut
gpost	postnominales Genitivattribut

4.18. morph_ellipsis

Typ: *Layer* – morph_ellipsis

Name:	morph_ellipsis
Dokumentation:	Annotation von Wortteilellipsen, z.B. *Anis- und Eppichsamen.* Informationen zum gesamten Konvertierungsworkflow gibt es hier: https://github.com/langbank/conversion-pipeline.
Beschreibung:	Unter linguistischen Annotationen werden u.a. pos-Tagging, Lemmatisierung und Markierung von fremdsprachlichem Material zusammengefasst.

Typ: Preparationstep – morph_ellipsis

Schritt:	1
Aufbereitung:	Linguistische Annotation
Aufbereitungsart:	manuell
Tool	Excel 2010
Format:	Excel 2010
Segmentierung der Annotation:	basiert auf ‘dipl’
Qualitätsprüfer:	Laura Perlitz
Datum	2018
Qualitätsprüfung:	manuelle Korrektur
Editor:	Laura Perlitz

Schritt:	2
Aufbereitung:	Konvertierung von Excel 2013 nach PAULA-XML und ANNIS
Aufbereitungsart:	automatisch
Tool:	Pepper, Konvertierung des Formats Excel nach ANNIS und PAULA mithilfe der folgenden Importer: SpreadsheetImporter. Export ins ANNIS- und PAULA-Format mit dem ANNISExporter und dem PAULAExporter (https://korpling.german.hu-berlin.de/saltnpepper/pepper/download/stable/Pepper_2019.06.11.zip)
Format:	PAULA-XML, ANNIS
Segmentierung der Annotation:	basiert auf 'dipl'
Qualitätsprüfer:	korpling
Datum:	2018
Qualitätsprüfung:	NA
Editor:	LAUDATIO, LangBank

Typ: Annotationlayer – morph_ellipsis

Annotationsart:	Spannenannotation über mehrere dipl-Einheiten
Beschreibung:	Markiert alle im Korpus enthaltenen koordinierten Komposita („strukturelle Destruktionen"; z.B.: *gelb≈ und Waſſerfucht*).

Typ: Annotationvalue – morph_ellipsis

Wert:	Wertbeschreibung:
morph_ellipsis	Wortteilellipse (morphologische Ellipse)

4.19. kaus1

Typ: Layer – kaus1

Name:	kaus1
Dokumentation:	Informationen zum gesamten Konvertierungsworkflow gibt es hier: https://github.com/langbank/conversion-pipeline.
Beschreibung:	Unter linguistischen Annotationen werden u.a. pos-Tagging,

	Lemmatisierung und Markierung von fremdsprachlichem Material zusammengefasst.

Typ: Preparationstep – kaus1

Schritt:	1
Aufbereitung:	Linguistische Annotation
Aufbereitungsart:	manuell
Tool	Excel 2013
Format:	Excel 2013
Segmentierung der Annotation:	basiert auf 'dipl'
Qualitätsprüfer:	Catharina Fischer
Datum	2019
Qualitätsprüfung:	manuell
Editor:	Catharina Fischer

Schritt:	2
Aufbereitung:	Konvertierung von Excel 2013 nach PAULA-XML und ANNIS
Aufbereitungsart:	automatisch
Tool:	Pepper, Konvertierung des Formats Excel nach ANNIS und PAULA mithilfe der folgenden Importer: SpreadsheetImporter. Export ins ANNIS- und PAULA-Format mit dem ANNISExporter und dem PAULAExporter (https://korpling.german.hu-berlin.de/saltnpepper/pepper/download/stable/Pepper_2019.06.11.zip)
Format:	PAULA-XML, ANNIS
Segmentierung der Annotation:	basiert auf 'dipl'
Qualitätsprüfer:	korpling
Datum:	2019
Qualitätsprüfung:	NA
Editor:	Korpling

Typ: Annotationlayer – kaus1

Annotationsart:	Spannenannotation über mehrere dipl-Einheiten
Beschreibung:	Erste Annotationsebene von Antezedens und Konsequens in einer kausalen Relation

Typ: Annotationvalue - kaus1

Wert:	Wertbeschreibung:
ant	Antezedens einer kausalen Relation
kons	Konsequens einer kausalen Relation

4.20. kaus2

Typ: Layer – kaus1

Name:	kaus2
Dokumentation:	Informationen zum gesamten Konvertierungsworkflow gibt es hier: https://github.com/langbank/conversion-pipeline.
Beschreibung:	Unter linguistischen Annotationen werden u.a. pos-Tagging, Lemmatisierung und Markierung von fremdsprachlichem Material zusammengefasst.

Typ: Preparationstep – kaus2

Schritt:	1
Aufbereitung:	Linguistische Annotation
Aufbereitungsart:	manuell
Tool	Excel 2013
Format:	Excel 2013
Segmentierung der Annotation:	basiert auf ‘dipl’
Qualitätsprüfer:	Catharina Fischer
Datum	2018
Qualitätsprüfung:	manuell
Editor:	Catharina Fischer

Schritt:	2

Aufbereitung:	Konvertierung von Excel 2013 nach PAULA-XML und ANNIS
Aufbereitungsart:	automatisch
Tool:	Pepper, Konvertierung des Formats Excel nach ANNIS und PAULA mithilfe der folgenden Importer: SpreadsheetImporter. Export ins ANNIS- und PAULA-Format mit dem ANNISExporter und dem PAULAExporter (https://korpling.german.hu-berlin.de/saltnpepper/pepper/download/stable/Pepper_2019.06.11.zip)
Format:	PAULA-XML, ANNIS
Segmentierung der Annotation:	basiert auf 'dipl'
Qualitätsprüfer:	korpling
Datum:	2019
Qualitätsprüfung:	NA
Editor:	Korpling

Typ: Annotationlayer – kaus2

Annotationsart:	Spannenannotation über mehrere dipl-Einheiten
Beschreibung:	Zweite Annotationsebene von Antezedens und Konsequens einer kausalen Relation

Typ: Annotationvalue - kaus2

Wert:	Wertbeschreibung:
ant	Antezedens einer kausalen Relation
kons	Konsequens einer kausalen Relation

4.21. kaus3

Typ: Layer – kaus3

Name:	kaus3
Dokumentation:	Informationen zum gesamten Konvertierungsworkflow gibt es

Beschreibung:	hier: https://github.com/langbank/conversion-pipeline. Unter linguistischen Annotationen werden u.a. pos-Tagging, Lemmatisierung und Markierung von fremdsprachlichem Material zusammengefasst.

Typ: Preparationstep – kaus3

Schritt:	1
Aufbereitung:	Linguistische Annotation
Aufbereitungsart:	manuell
Tool	Excel 2013
Format:	Excel 2013
Segmentierung der Annotation:	basiert auf 'dipl'
Qualitätsprüfer:	Catharina Fischer
Datum	2019
Qualitätsprüfung:	manuell
Editor:	Catharina Fischer

Schritt:	2
Aufbereitung:	Konvertierung von Excel 2013 nach PAULA-XML und ANNIS
Aufbereitungsart:	automatisch
Tool:	Pepper, Konvertierung des Formats Excel nach ANNIS und PAULA mithilfe der folgenden Importer: SpreadsheetImporter. Export ins ANNIS- und PAULA-Format mit dem ANNISExporter und dem PAULAExporter (https://korpling.german.hu-berlin.de/saltnpepper/pepper/download/stable/Pepper_2019.06.11.zip)
Format:	PAULA-XML, ANNIS
Segmentierung der Annotation:	basiert auf 'dipl'
Qualitätsprüfer:	korpling
Datum:	2019
Qualitätsprüfung:	NA
Editor:	Korpling

Typ: Annotationlayer – kaus3

Annotationsart:	Spannenannotation über mehrere dipl-Einheiten
Beschreibung:	Dritte Annotationsebene von Antezedens und Konsequens einer kausalen Relation

Typ: Annotationvalue - kaus3

Wert:	Wertbeschreibung:
ant	Antezedens einer kausalen Relation
kons	Konsequens einer kausalen Relation

4.22. kaus4

Typ: Layer – kaus4

Name:	kaus4
Dokumentation:	Informationen zum gesamten Konvertierungsworkflow gibt es hier: https://github.com/langbank/conversion-pipeline.
Beschreibung:	Unter linguistischen Annotationen werden u.a. pos-Tagging, Lemmatisierung und Markierung von fremdsprachlichem Material zusammengefasst.

Typ: Preparationstep – kaus4

Schritt:	1
Aufbereitung:	Linguistische Annotation
Aufbereitungsart:	manuell
Tool	Excel 2013
Format:	Excel 2013
Segmentierung der Annotation:	basiert auf ‘dipl’
Qualitätsprüfer:	Catharina Fischer
Datum	2019
Qualitätsprüfung:	manuell
Editor:	Catharina Fischer

Schritt:	2

Aufbereitung:	Konvertierung von Excel 2013 nach PAULA-XML und ANNIS
Aufbereitungsart:	automatisch
Tool:	Pepper, Konvertierung des Formats Excel nach ANNIS und PAULA mithilfe der folgenden Importer: SpreadsheetImporter. Export ins ANNIS- und PAULA-Format mit dem ANNISExporter und dem PAULAExporter (https://korpling.german.hu-berlin.de/saltnpepper/pepper/download/stable/Pepper_2019.06.11.zip)
Format:	PAULA-XML, ANNIS
Segmentierung der Annotation:	basiert auf 'dipl'
Qualitätsprüfer:	korpling
Datum:	2019
Qualitätsprüfung:	NA
Editor:	Korpling

Typ: Annotationlayer – kaus4

Annotationsart:	Spannenannotation über mehrere dipl-Einheiten
Beschreibung:	Dritte Annotationsebene von Antezedens und Konsequens einer kausalen Relation

Typ: Annotationvalue - kaus4

Wert:	Wertbeschreibung:
ant	Antezedens einer kausalen Relation
kons	Konsequens einer kausalen Relation

4.23. kon1

Typ: Layer – kon1

Name:	kon1
Dokumentation:	Informationen zum gesamten Konvertierungsworkflow gibt es hier: https://github.com/langbank/conversion-pipeline.
Beschreibung:	Unter linguistischen Annotationen werden u.a. pos-Tagging,

	Lemmatisierung und Markierung von fremdsprachlichem Material zusammengefasst.

Typ: Preparationstep – kon1

Schritt:	1
Aufbereitung:	Linguistische Annotation
Aufbereitungsart:	manuell
Tool	Excel 2013
Format:	Excel 2013
Segmentierung der Annotation:	basiert auf ‘norm’
Qualitätsprüfer:	Catharina Fischer
Datum	2019
Qualitätsprüfung:	manuell
Editor:	Catharina Fischer
Schritt:	2
Aufbereitung:	Konvertierung von Excel 2013 nach PAULA-XML und ANNIS
Aufbereitungsart:	automatisch
Tool:	Pepper, Konvertierung des Formats Excel nach ANNIS und PAULA mithilfe der folgenden Importer: SpreadsheetImporter. Export ins ANNIS- und PAULA-Format mit dem ANNISExporter und dem PAULAExporter (https://korpling.german.hu-berlin.de/saltnpepper/pepper/download/stable/Pepper_2019.06.11.zip)
Format:	PAULA-XML, ANNIS
Segmentierung der Annotation:	basiert auf ‘norm’
Qualitätsprüfer:	korpling
Datum:	2019
Qualitätsprüfung:	NA
Editor:	Korpling

Typ: Annotationlayer – kon1

Annotationsart:	Annotationswert wird je norm-Einheit zugewiesen
Beschreibung:	Erste Annotationsebene der Art des Konnektors in kausalen Relationen (basierend auf „kaus“).

Typ: Annotationvalue - kon1

Wert:	Wertbeschreibung:
konk	Kausaler Konnektor nach dem Handbuch deutscher Konnektoren (HdK)
kone	Konnektor (nicht kausal nach dem HdK)
koni	kein Konnketor
kona	Annotation von *und* oder *auch* in Kombination mit einem weiteren Konnektor am Anfang einer Relation

4.24. kon2

Typ: Layer – kon2

Name:	kon2
Dokumentation:	Informationen zum gesamten Konvertierungsworkflow gibt es hier: https://github.com/langbank/conversion-pipeline.
Beschreibung:	Unter linguistischen Annotationen werden u.a. pos-Tagging, Lemmatisierung und Markierung von fremdsprachlichem Material zusammengefasst.

Typ: Preparationstep – kon2

Schritt:	1
Aufbereitung:	Linguistische Annotation
Aufbereitungsart:	manuell
Tool	Excel 2013
Format:	Excel 2013
Segmentierung der Annotation:	basiert auf ‘norm’
Qualitätsprüfer:	Catharina Fischer
Datum	2019
Qualitätsprüfung:	manuell
Editor:	Catharina Fischer
Schritt:	2
Aufbereitung:	Konvertierung von Excel 2013 nach PAULA-XML und ANNIS
Aufbereitungsart:	automatisch
Tool:	Pepper, Konvertierung des Formats Excel nach ANNIS und PAULA mithilfe der folgenden Importer: SpreadsheetImporter.

	Export ins ANNIS- und PAULA-Format mit dem ANNISExporter und dem PAULAExporter (https://korpling.german.hu-berlin.de/saltnpepper/pepper/download/stable/Pepper_2019.06.11.zip)
Format:	PAULA-XML, ANNIS
Segmentierung der Annotation:	basiert auf 'norm'
Qualitätsprüfer:	korpling
Datum:	2019
Qualitätsprüfung:	NA
Editor:	Korpling

Typ: Annotationlayer – kon2

Annotationsart:	Annotationswert wird je norm-Einheit zugewiesen
Beschreibung:	Zweite Annotationsebene der Art des Konnektors in kausalen Relationen (basierend auf „kaus").

Typ: Annotationvalue - kon2

Wert:	Wertbeschreibung:
konk	Kausaler Konnektor nach dem Handbuch deutscher Konnektoren (HdK)
kone	Konnektor (nicht kausal nach dem HdK)
koni	kein Konnketor
kona	Annotation von *und* oder *auch* in Kombination mit einem weiteren Konnektor am Anfang einer Relation

4.25. lin1

Typ: Layer – lin1

Name:	lin1
Dokumentation:	Informationen zum gesamten Konvertierungsworkflow gibt es hier: https://github.com/langbank/conversion-pipeline.
Beschreibung:	Unter linguistischen Annotationen werden u.a. pos-Tagging, Lemmatisierung und Markierung von fremdsprachlichem Material zusammengefasst.

Typ: Preparationstep – lin1

Schritt:	1
Aufbereitung:	Linguistische Annotation
Aufbereitungsart:	manuell
Tool	Excel 2013
Format:	Excel 2013
Segmentierung der Annotation:	basiert auf ‘dipl’
Qualitätsprüfer:	Catharina Fischer
Datum	2019
Qualitätsprüfung:	manuell
Editor:	Catharina Fischer

Schritt:	2
Aufbereitung:	Konvertierung von Excel 2013 nach PAULA-XML und ANNIS
Aufbereitungsart:	automatisch
Tool:	Pepper, Konvertierung des Formats Excel nach ANNIS und PAULA mithilfe der folgenden Importer: SpreadsheetImporter. Export ins ANNIS- und PAULA-Format mit dem ANNISExporter und dem PAULAExporter (https://korpling.german.hu-berlin.de/saltnpepper/pepper/download/stable/Pepper_2019.06.11.zip)
Format:	PAULA-XML, ANNIS
Segmentierung der Annotation:	basiert auf ‘dipl’
Qualitätsprüfer:	korpling
Datum:	2019
Qualitätsprüfung:	NA
Korpling	Korpling

Typ: Annotationlayer – lin1

Annotationsart:	Spannenannotation über mehrere dipl-Einheiten
Beschreibung:	Erste Annotationsebene der Relation der Konnekte (Abfolge von Antezedens und Konseques basierend auf „kaus“)

Typ: Annotationvalue - lin1

Wert:	Wertbeschreibung:
ik	Konsequens folgt Antezeden
ika	Antezedens folgt Konsequens
ikinsert	Konsequens ist in Antezedens eingebettet
ikainsert	Antezedens ist in Konsequens eingebettet

4.26. lin2

Typ: Layer – lin2

Name:	lin2
Dokumentation:	Informationen zum gesamten Konvertierungsworkflow gibt es hier: https://github.com/langbank/conversion-pipeline.
Beschreibung:	Unter linguistischen Annotationen werden u.a. pos-Tagging, Lemmatisierung und Markierung von fremdsprachlichem Material zusammengefasst.

Typ: Preparationstep – lin2

Schritt:	1
Aufbereitung:	Linguistische Annotation
Aufbereitungsart:	manuell
Tool	Excel 2013
Format:	Excel 2013
Segmentierung der Annotation:	basiert auf ‘dipl’
Qualitätsprüfer:	Catharina Fischer
Datum	2019
Qualitätsprüfung:	manuell
Editor:	Catharina Fischer

Schritt:	2
Aufbereitung:	Konvertierung von Excel 2013 nach PAULA-XML und ANNIS
Aufbereitungsart:	automatisch
Tool:	Pepper, Konvertierung des Formats Excel nach ANNIS und

	PAULA mithilfe der folgenden Importer: SpreadsheetImporter. Export ins ANNIS- und PAULA-Format mit dem ANNISExporter und dem PAULAExporter (https://korpling.german.hu-berlin.de/saltnpepper/pepper/download/stable/Pepper_2019.06.11.zip)
Format:	PAULA-XML, ANNIS
Segmentierung der Annotation:	basiert auf 'dipl'
Qualitätsprüfer:	korpling
Datum:	2019
Qualitätsprüfung:	NA
Editor:	Korpling

Typ: Annotationlayer – lin2

Annotationsart:	Spannenannotation über mehrere dipl-Einheiten
Beschreibung:	Zweite Annotationsebene der Relation der Konnekte (Abfolge von Antezedens und Konseques basierend auf „kaus“)

Typ: Annotationvalue - lin2

Wert:	Wertbeschreibung:
ik	Konsequens folgt Antezeden
ika	Antezedens folgt Konsequens
ikinsert	Konsequens ist in Antezedens eingebettet
ikainsert	Antezedens ist in Konsequens eingebettet

4.27. lin3

Typ: Layer – lin3

Name:	lin3
Dokumentation:	Informationen zum gesamten Konvertierungsworkflow gibt es hier: https://github.com/langbank/conversion-pipeline.
Beschreibung:	Unter linguistischen Annotationen werden u.a. pos-Tagging, Lemmatisierung und Markierung von fremdsprachlichem Material zusammengefasst.

Typ: Preparationstep – lin3

Schritt:	1
Aufbereitung:	Linguistische Annotation
Aufbereitungsart:	manuell
Tool	Excel 2013
Format:	Excel 2013
Segmentierung der Annotation:	basiert auf ‘dipl’
Qualitätsprüfer:	Catharina Fischer
Datum	2019
Qualitätsprüfung:	manuell
Editor:	Catharina Fischer

Schritt:	2
Aufbereitung:	Konvertierung von Excel 2013 nach PAULA-XML und ANNIS
Aufbereitungsart:	automatisch
Tool:	Pepper, Konvertierung des Formats Excel nach ANNIS und PAULA mithilfe der folgenden Importer: SpreadsheetImporter. Export ins ANNIS- und PAULA-Format mit dem ANNISExporter und dem PAULAExporter (https://korpling.german.hu-berlin.de/saltnpepper/pepper/download/stable/Pepper_2019.06.11.zip)
Format:	PAULA-XML, ANNIS
Segmentierung der Annotation:	basiert auf ‘dipl’
Qualitätsprüfer:	korpling
Datum:	2019
Qualitätsprüfung:	NA
Editor:	Korpling

Typ: Annotationlayer – lin3

Annotationsart:	Spannenannotation über mehrere dipl-Einheiten
Beschreibung:	Dritte Annotationsebene der Relation der Konnekte (Abfolge von Antezedens und Konseques basierend auf „kaus“)

Typ: Annotationvalue - lin3

Wert:	Wertbeschreibung:
ik	Konsequens folgt Antezeden
ika	Antezedens folgt Konsequens
ikinsert	Konsequens ist in Antezedens eingebettet
ikainsert	Antezedens ist in Konsequens eingebettet

4.28. lin4

Typ: Layer – lin4

Name:	lin4
Dokumentation:	Informationen zum gesamten Konvertierungsworkflow gibt es hier: https://github.com/langbank/conversion-pipeline.
Beschreibung:	Unter linguistischen Annotationen werden u.a. pos-Tagging, Lemmatisierung und Markierung von fremdsprachlichem Material zusammengefasst.

Typ: Preparationstep – lin4

Schritt:	1
Aufbereitung:	Linguistische Annotation
Aufbereitungsart:	manuell
Tool	Excel 2013
Format:	Excel 2013
Segmentierung der Annotation:	basiert auf ‘dipl’
Qualitätsprüfer:	Catharina Fischer
Datum	2019
Qualitätsprüfung:	manuell
Editor:	Catharina Fischer

Schritt:	2
Aufbereitung:	Konvertierung von Excel 2013 nach PAULA-XML und ANNIS

Aufbereitungsart:	automatisch
Tool:	Pepper, Konvertierung des Formats Excel nach ANNIS und PAULA mithilfe der folgenden Importer: SpreadsheetImporter. Export ins ANNIS- und PAULA-Format mit dem ANNISExporter und dem PAULAExporter (https://korpling.german.hu-berlin.de/saltnpepper/pepper/download/stable/Pepper_2019.06.11.zip)
Format:	PAULA-XML, ANNIS
Segmentierung der Annotation:	basiert auf 'dipl'
Qualitätsprüfer:	korpling
Datum:	2019
Qualitätsprüfung:	NA
Editor:	Korpling

Typ: Annotationlayer – lin4

Annotationsart:	Spannenannotation über mehrere dipl-Einheiten
Beschreibung:	Vierte Annotationsebene der Relation der Konnekte (Abfolge von Antezedens und Konseques basierend auf „kaus“)

Typ: Annotationvalue - lin4

Wert:	Wertbeschreibung:
ik	Konsequens folgt Antezeden
ika	Antezedens folgt Konsequens
ikinsert	Konsequens ist in Antezedens eingebettet
ikainsert	Antezedens ist in Konsequens eingebettet

4.29. hypo

Typ: Layer – hypo

Name:	hypo
Dokumentation:	Informationen zum gesamten Konvertierungsworkflow gibt es hier: https://github.com/langbank/conversion-pipeline.
Beschreibung:	Unter linguistischen Annotationen werden u.a. pos-Tagging, Lemmatisierung und Markierung von fremdsprachlichem

Material zusammengefasst.

Typ: Preparationstep – hypo

Schritt:	1
Aufbereitung:	Linguistische Annotation
Aufbereitungsart:	manuell
Tool	Excel 2013
Format:	Excel 2013
Segmentierung der Annotation:	basiert auf ‘dipl’
Qualitätsprüfer:	Catharina Fischer
Datum	2019
Qualitätsprüfung:	manuell
Editor:	Catharina Fischer

Schritt:	2
Aufbereitung:	Konvertierung von Excel 2013 nach PAULA-XML und ANNIS
Aufbereitungsart:	automatisch
Tool:	Pepper, Konvertierung des Formats Excel nach ANNIS und PAULA mithilfe der folgenden Importer: SpreadsheetImporter. Export ins ANNIS- und PAULA-Format mit dem ANNISExporter und dem PAULAExporter (https://korpling.german.hu-berlin.de/saltnpepper/pepper/download/stable/Pepper_2019.06.11.zip)
Format:	PAULA-XML, ANNIS
Segmentierung der Annotation:	basiert auf ‘dipl’
Qualitätsprüfer:	korpling
Datum:	2019
Qualitätsprüfung:	NA
Editor:	Korpling

Typ: Annotationlayer – hypo

Annotationsart:	Spannenannotation über mehrere dipl-Einheiten
Beschreibung:	Verknüpfung der kausalen Relationen (basiered auf „kaus“)

Typ: Annotationvalue - hypo

Wert:	Wertbeschreibung:
hypo	Markierung von ineinander eingebetteteten kausalen Relationen

4.30. M

Typ: Layer – M

Name:	M
Dokumentation:	Informationen zum gesamten Konvertierungsworkflow gibt es hier: https://github.com/langbank/conversion-pipeline.
Beschreibung:	Unter linguistischen Annotationen werden u.a. pos-Tagging, Lemmatisierung und Markierung von fremdsprachlichem Material zusammengefasst.

Typ: Preparationstep – M

Schritt:	1
Aufbereitung:	Linguistische Annotation
Aufbereitungsart:	manuell
Tool	Excel 2013
Format:	Excel 2013
Segmentierung der Annotation:	basiert auf ‘norm’
Qualitätsprüfer:	Gohar Schnelle, Carolin Odebrecht
Datum	2019
Qualitätsprüfung:	manuell
Editor:	Gohar Schnelle, Carolin Odebrecht

Schritt:	2
Aufbereitung:	Konvertierung von Excel 2013 nach PAULA-XML und ANNIS
Aufbereitungsart:	automatisch
Tool:	Pepper, Konvertierung des Formats Excel nach ANNIS und PAULA mithilfe der folgenden Importer: SpreadsheetImporter. Export ins ANNIS- und PAULA-Format mit dem ANNISExporter und dem PAULAExporter (https://korpling.german.hu-berlin.de/saltnpepper/pepper/download/stable/Pepper_2019.06.11

	.zip)
Format:	PAULA-XML, ANNIS
Segmentierung der Annotation:	basiert auf ‘norm’
Qualitätsprüfer:	korpling
Datum:	2019
Qualitätsprüfung:	NA
Editor:	Korpling

Typ: Annotationlayer – M

Annotationsart:	Spannenannotation über mehrere norm-Einheiten
Beschreibung:	Annotation von „Menstruation“

Typ: Annotationvalue - M

Wert:	Wertbeschreibung:
M	Annotation der minimalen Nominalphrase (exklusive Präpositionen oder höher eingebetteten Phrasen)

4.31. gram

Typ: Layer – gram

Name:	gram
Dokumentation:	Informationen zum gesamten Konvertierungsworkflow gibt es hier: https://github.com/langbank/conversion-pipeline.
Beschreibung:	Unter linguistischen Annotationen werden u.a. pos-Tagging, Lemmatisierung und Markierung von fremdsprachlichem Material zusammengefasst.

Typ: Preparationstep – gram

Schritt:	1
Aufbereitung:	Linguistische Annotation
Aufbereitungsart:	manuell
Tool	Excel 2013
Format:	Excel 2013
Segmentierung der Annotation:	basiert auf ‘norm’

Qualitätsprüfer:	Gohar Schnelle, Carolin Odebrecht
Datum	2019
Qualitätsprüfung:	manuell
Editor:	Gohar Schnelle, Carolin Odebrecht
Schritt:	2
Aufbereitung:	Konvertierung von Excel 2013 nach PAULA-XML und ANNIS
Aufbereitungsart:	automatisch
Tool:	Pepper, Konvertierung des Formats Excel nach ANNIS und PAULA mithilfe der folgenden Importer: SpreadsheetImporter. Export ins ANNIS- und PAULA-Format mit dem ANNISExporter und dem PAULAExporter (https://korpling.german.hu-berlin.de/saltnpepper/pepper/download/stable/Pepper_2019.06.11.zip)
Format:	PAULA-XML, ANNIS
Segmentierung der Annotation:	basiert auf 'norm'
Qualitätsprüfer:	korpling
Datum:	2019
Qualitätsprüfung:	NA
Editor:	Korpling

Typ: Annotationlayer – gram

Annotationsart:	Spannenannotation über mehrere norm-Einheiten
Beschreibung:	Zurodnung grammatikalischer Einheiten der Nominalphrase, sowie Klassifizierung der morphologischen Komplexität des nominalen Kerns

Typ: Annotationvalue - gram

Wert:	Wertbeschreibung:
det	Determinierer der Nominalphrase mit Referenz auf "Menstruation"
adjAtt	Adjektiv Attribut der Nominalphrase mit Referenz auf "Menstruation"
genAtt	Genitiv Attribut der Nominalphrase mit Referenz auf "Menstruation"

objAtt	Objekattribut der Nominalphrase mit Referenz auf “Menstruation”
relAtt	Relativsatzattribut der Nominalphrase mit Referenz auf “Menstruation”
nks	Nomenkern Simplex
nkc	Nomenkern Kompositum
nkd	Nomenkern Derivation

4.32. technicalterm

Typ: Layer – technicalterm

Name:	technicalterm
Dokumentation:	Informationen zum gesamten Konvertierungsworkflow gibt es hier: https://github.com/langbank/conversion-pipeline.
Beschreibung:	Unter linguistischen Annotationen werden u.a. pos-Tagging, Lemmatisierung und Markierung von fremdsprachlichem Material zusammengefasst.

Typ: Preparationstep – technicalterm

Schritt:	1
Aufbereitung:	Linguistische Annotation
Aufbereitungsart:	manuell
Tool	Excel 2013
Format:	Excel 2013
Segmentierung der Annotation:	basiert auf ‘norm’
Qualitätsprüfer:	Gohar Schnelle, Carolin Odebrecht
Datum	2019
Qualitätsprüfung:	manuell
Editor:	Gohar Schnelle, Carolin Odebrecht

Schritt:	2
Aufbereitung:	Konvertierung von Excel 2013 nach PAULA-XML und ANNIS
Aufbereitungsart:	automatisch
Tool:	Pepper, Konvertierung des Formats Excel nach ANNIS und PAULA mithilfe der folgenden Importer: SpreadsheetImporter.

	Export ins ANNIS- und PAULA-Format mit dem ANNISExporter und dem PAULAExporter (https://korpling.german.hu-berlin.de/saltnpepper/pepper/download/stable/Pepper_2019.06.11.zip)
Format:	PAULA-XML, ANNIS
Segmentierung der Annotation:	basiert auf 'norm'
Qualitätsprüfer:	korpling
Datum:	2019
Qualitätsprüfung:	NA
Editor:	Korpling

Typ: Annotationlayer – technicalterm

Annotationsart:	Spannenannotation über mehrere norm-Einheiten
Beschreibung:	Semiotische Variation: Art der Beziehung zwischen der Wahl des vernakulären Zeiches, seiner grundsätzlichen Bedeutung und den Eigenschaften des medizinischen Konzepts auf das es referiert.

Typ: Annotationvalue - technicalterm

Wert:	Wertbeschreibung:
para	Paraphrase
met	Metapher
loanmed	Medizinisches Lehnwort aus klassischen Sprachen
her	Ererbtes Vokabular, das zur Modifikation von Referenzen verwendet wird (nur für Attribute annotiert)
herspec	Ererbtes Vokabular, das zur Modifikation von Referenzen verwendet wird (nur für Attribute annotiert)

5. Annotationsebenen – Strukturelle Annotation

5.1. lb

Typ: *Layer* – lb

Name:	lb
Dokumentation:	Informationen zum gesamten Konvertierungsworkflow gibt es

	hier: https://github.com/langbank/conversion-pipeline.
Beschreibung:	Unter strukturellen Annotationen werden alle Annotationen zusammengefasst, die Aussagen über die graphische Strukturierung der Textgrundlage liefern. Dazu zählen Kennzeichnungen der Zeilen- und Seitenumbrüche, Seitenzahlen, unkenntliche/nicht lesbare Zeichen und Marginalien sowie Fußnoten. Strukturelle Annotationen werden auf Grundlage der diplomatischen Ebene vergeben (**dipl**-Ebene). So ist die Tokenisierung in der **dipl**-Ebene Vorlage für die weiteren, strukturellen Annotationen.

Typ: Preparationstep – lb

Die lb-Ebene wird für gewöhnlich folgendermaßen erstellt (für die Aufbereitungsschritte mit ocr in den Dokumenten der Version 7 siehe hier):

Schritt:	1a
Aufbereitung:	Strukturelle Annotation
Aufbereitungsart:	manuell
Tool	Excel 2013
Format:	Excel 2013
Segmentierung der Annotation:	basiert auf ‘dipl’
Qualitätsprüfer:	korpling
Datum:	2019
Qualitätsprüfung:	manuelle Korrektur
Editor:	Studierende der Humboldt-Universität zu Berlin und der Georg-August-Universität Göttingen, korpling
Schritt:	2a
Aufbereitung:	Konvertierung von Excel 2013 nach PAULA-XML und ANNIS
Aufbereitungsart:	automatisch
Tool:	Pepper, Konvertierung des Formats Excel nach ANNIS und PAULA mithilfe der folgenden Importer: SpreadsheetImporter. Export ins ANNIS- und PAULA-Format mit dem ANNISExporter und dem PAULAExporter (https://korpling.german.hu-berlin.de/saltnpepper/pepper/download/stable/Pepper_2019.06.11.zip)
Format:	PAULA-XML, ANNIS
Segmentierung der Annotation:	basiert auf ‘dipl’
Qualitätsprüfer:	korpling

Datum:	2019
Qualitätsprüfung:	manuelle Korrektur
Editor:	Korpling

Typ: Annotationlayer – lb

Annotationsart:	Spannenannotation über mehrere dipl-Einheiten
Beschreibung:	Markierung einer ganzen Zeile.

Typ: Annotationvalue – lb

Wert:	Wertbeschreibung:
lb	Markierung pro ganzer Zeile.

5.2. quotemark

Typ: *Layer* – quotemarke

Name:	quotemarke
Dokumentation:	Informationen zum gesamten Konvertierungsworkflow gibt es hier: https://github.com/langbank/conversion-pipeline.
Beschreibung:	Unter strukturellen Annotationen werden alle Annotationen zusammengefasst, die Aussagen über die graphische Strukturierung der Textgrundlage liefern. Dazu zählen Kennzeichnungen der Zeilen- und Seitenumbrüche, Seitenzahlen, unkenntliche/nicht lesbare Zeichen und Marginalien sowie Fußnoten. Strukturelle Annotationen werden auf Grundlage der diplomatischen Ebene vergeben (**dipl**-Ebene). So ist die Tokenisierung in der **dipl**-Ebene Vorlage für die weiteren, strukturellen Annotationen.

Typ: Preparationstep – quotemarke

Schritt:	1

Aufbereitung:	Linguistische Annotation
Aufbereitungsart:	manuell
Tool	Excel 2013
Format:	Excel 2013
Segmentierung der Annotation:	basiert auf ‘dipl’
Qualitätsprüfer:	Catharina Fischer
Datum	2019
Qualitätsprüfung:	manuell
Editor:	Catharina Fischer

Schritt:	2
Aufbereitung:	Konvertierung von Excel 2013 nach PAULA-XML und ANNIS
Aufbereitungsart:	automatisch
Tool:	Pepper, Konvertierung des Formats Excel nach ANNIS und PAULA mithilfe der folgenden Importer: SpreadsheetImporter. Export ins ANNIS- und PAULA-Format mit dem ANNISExporter und dem PAULAExporter (https://korpling.german.hu-berlin.de/saltnpepper/pepper/download/stable/Pepper_2019.06.11.zip)
Format:	PAULA-XML, ANNIS
Segmentierung der Annotation:	basiert auf ‘dipl’
Qualitätsprüfer:	korpling
Datum:	2019
Qualitätsprüfung:	NA
Editor:	Korpling

Typ: Annotationlayer – quotemark

Annotationsart:	Spannenannotation über mehrere dipl-Einheiten
Beschreibung:	Markierung von Anführungszeichen (außerhalb des Fließtextes)

Typ: Annotationvalue – quotemark

Wert:	Wertbeschreibung:
qm	Markierung pro Zeile mit Anführungszeichen

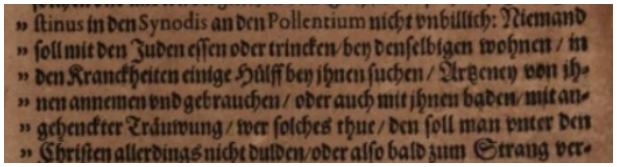


Bild: Beschreibung der Gräser nebst ihren Abbildungen nach der Natur / Johann Christian Daniel Schrebers (Leipzig, 1769) Zentralbibliothek Zürich: https://www.e-rara.ch/zuz/content/pageview/16384579

5.3. column

Typ: *Layer* – column

Name:	column
Dokumentation:	Informationen zum gesamten Konvertierungsworkflow gibt es hier: https://github.com/langbank/conversion-pipeline.
Beschreibung:	Unter strukturellen Annotationen werden alle Annotationen zusammengefasst, die Aussagen über die graphische Strukturierung der Textgrundlage liefern. Dazu zählen Kennzeichnungen der Zeilen- und Seitenumbrüche, Seitenzahlen, unkenntliche/nicht lesbare Zeichen und Marginalien sowie Fußnoten. Strukturelle Annotationen werden auf Grundlage der diplomatischen Ebene vergeben (**dipl**-Ebene). So ist die Tokenisierung in der **dipl**-Ebene Vorlage für die weiteren, strukturellen Annotationen.

Typ: Preparationstep – column

Schritt:	1
Aufbereitung:	Strukturelle Annotation
Aufbereitungsart:	manuell
Tool	Excel 2013
Format:	Excel 2013
Segmentierung der Annotation:	basiert auf ‘dipl’
Qualitätsprüfer:	korpling
Datum:	2019
Qualitätsprüfung:	manuelle Korrektur
Editor:	Studierende der Humboldt-Universität zu Berlin, korpling

Schritt:	2
Aufbereitung:	Konvertierung von Excel 2013 nach PAULA-XML und ANNIS
Aufbereitungsart:	automatisch
Tool:	Pepper, Konvertierung des Formats Excel nach ANNIS und PAULA mithilfe der folgenden Importer: SpreadsheetImporter. Export ins ANNIS- und PAULA-Format mit dem ANNISExporter und dem PAULAExporter (https://korpling.german.hu-berlin.de/saltnpepper/pepper/download/stable/Pepper_2019.06.11.zip)
Format:	PAULA-XML, ANNIS
Segmentierung der Annotation:	basiert auf 'dipl'
Qualitätsprüfer:	korpling
Datum:	2019
Qualitätsprüfung:	manuelle Korrektur
Editor:	Korpling

Typ: Annotationlayer – column

Annotationsart:	Spannenannotation über mehrere dipl-Einheiten
Beschreibung:	Markierung einer ganzen Spalte

Typ: Annotationvalue – column

Wert:	Wertbeschreibung:
l	dipl-Einheiten, die zusammen in der linken Spalte stehen
r	dipl-Einheiten, die zusammen in der rechten Spalte stehen
c	bei dreispaltigem Text: dipl-Einheiten, die zusammen in der mittleren Spalte stehen (center)

5.4. pb

Typ: *Layer* – pb

Name:	pb

Dokumentation:	Informationen zum gesamten Konvertierungsworkflow gibt es hier: https://github.com/langbank/conversion-pipeline.
Beschreibung:	Unter strukturellen Annotationen werden alle Annotationen zusammengefasst, die Aussagen über die graphische Strukturierung der Textgrundlage liefern. Dazu zählen Kennzeichnungen der Zeilen- und Seitenumbrüche, Seitenzahlen, unkenntliche/nicht lesbare Zeichen und Marginalien sowie Fußnoten. Strukturelle Annotationen werden auf Grundlage der diplomatischen Ebene vergeben (**dipl**-Ebene). So ist die Tokenisierung in der **dipl**-Ebene Vorlage für die weiteren, strukturellen Annotationen.

Typ: Preparationstep – pb

Die pb-Ebene wird für gewöhnlich folgendermaßen erstellt (für die Aufbereitungsschritte mit ocr in den Dokumenten der Version 7 siehe hier):

Schritt:	1a
Aufbereitung:	Strukturelle Annotation
Aufbereitungsart:	manuell
Tool	Excel 2013
Format:	Excel 2013
Segmentierung der Annotation:	basiert auf 'dipl'
Qualitätsprüfer:	korpling
Datum:	2018
Qualitätsprüfung:	manuelle Korrektur
Editor:	Studierende der Humboldt-Universität zu Berlin und der Georg-August-Universität Göttingen, korpling
Schritt:	2a
Aufbereitung:	Konvertierung von Excel 2013 nach PAULA-XML und ANNIS
Aufbereitungsart:	automatisch
Tool:	Pepper, Konvertierung des Formats Excel nach ANNIS und PAULA mithilfe der folgenden Importer: SpreadsheetImporter. Export ins ANNIS- und PAULA-Format mit dem ANNISExporter und dem PAULAExporter (https://korpling.german.hu-berlin.de/saltnpepper/pepper/download/stable/Pepper_2019.06.11.zip)
Format:	PAULA-XML, ANNIS
Segmentierung der Annotation:	basiert auf 'dipl'

Qualitätsprüfer:	korpling
Datum:	2018
Qualitätsprüfung:	manuelle Korrektur
Editor:	Korpling

Typ: Annotationlayer – pb

Annotationsart:	Spannenannotation über mehrere dipl-Einheiten
Beschreibung:	Markierung einer ganzen Seite.

Typ: Annotationvalue – pb

Wert:	Wertbeschreibung:
pb	Markierung pro ganzer Seite.

5.5. pb_n

Typ: *Layer* – pb_n

Name:	pb_n
Dokumentation:	Informationen zum gesamten Konvertierungsworkflow gibt es hier: https://github.com/langbank/conversion-pipeline.
Beschreibung:	Unter strukturellen Annotationen werden alle Annotationen zusammengefasst, die Aussagen über die graphische Strukturierung der Textgrundlage liefern. Dazu zählen Kennzeichnungen der Zeilen- und Seitenumbrüche, Seitenzahlen, unkenntliche/nicht lesbare Zeichen und Marginalien sowie Fußnoten. Strukturelle Annotationen werden auf Grundlage der diplomatischen Ebene vergeben (**dipl**-Ebene). So ist die Tokenisierung in der **dipl**-Ebene Vorlage für die weiteren, strukturellen Annotationen.

Typ: Preparationstep – pb_n

Schritt:	1

Aufbereitung:	Strukturelle Annotation
Aufbereitungsart:	manuell
Tool	Excel 2013
Format:	Excel 2013
Segmentierung der Annotation:	basiert auf ‘dipl’
Qualitätsprüfer:	korpling
Datum:	2019
Qualitätsprüfung:	manuelle Korrektur
Editor:	Studierende der Humboldt-Universität zu Berlin und der Georg-August-Universität Göttingen, korpling

Schritt:	2
Aufbereitung:	Konvertierung von Excel 2013 nach PAULA-XML und ANNIS
Aufbereitungsart:	automatisch
Tool:	Pepper, Konvertierung des Formats Excel nach ANNIS und PAULA mithilfe der folgenden Importer: SpreadsheetImporter. Export ins ANNIS- und PAULA-Format mit dem ANNISExporter und dem PAULAExporter (https://korpling.german.hu-berlin.de/saltnpepper/pepper/download/stable/Pepper_2019.06.11.zip)
Format:	PAULA-XML, ANNIS
Segmentierung der Annotation:	basiert auf ‘dipl’
Qualitätsprüfer:	korpling
Datum:	2019
Qualitätsprüfung:	NA
Editor:	Korpling

Typ: Annotationlayer – pb_n

Annotationsart:	Spannenannotation über mehrere dipl-Einheiten
Beschreibung:	Markierung der Nummerierung einer Seite.

Typ: Annotationvalue – pb_n

Wert:	Wertbeschreibung:

numerischer Wert	Seitenzahl als lateinische oder römische Zahlenangabe (wird aus dem Faksimilé übernommen); ist keine Seitenangabe vorhanden, wird auch keine eingetragen

5.6. unclear

Typ: *Layer* - unclear

Name:	unclear
Dokumentation:	Informationen zum gesamten Konvertierungsworkflow gibt es hier: https://github.com/langbank/conversion-pipeline.
Beschreibung:	Unter strukturellen Annotationen werden alle Annotationen zusammengefasst, die Aussagen über die graphische Strukturierung der Textgrundlage liefern. Dazu zählen Kennzeichnungen der Zeilen- und Seitenumbrüche, Seitenzahlen, unkenntliche/nicht lesbare Zeichen und Marginalien sowie Fußnoten. Strukturelle Annotationen werden auf Grundlage der diplomatischen Ebene vergeben (dipl-Ebene). So ist die Tokenisierung in der dipl-Ebene Vorlage für die weiteren, strukturellen Annotationen.

Typ:*Preparationstep* – unclear

Schritt:	1
Aufbereitung:	Strukturelle Annotation
Aufbereitungsart:	manuell
Tool:	Excel 2013
Format:	Excel 2013
Segmentierung der Annotation:	basiert auf ‘dipl’
Qualitätsprüfer:	LAUDATIO
Datum	2019
Qualitätsprüfung:	Manuelle Konsistenzprüfung. Der Layer reason, der eine Begründung für unleserliche Elemente beschreiben sollte, wurde entfernt.
Editor:	Studierende der Humboldt-Universität zu Berlin und der Georg-August-Universität Göttingen

Schritt:	2
Aufbereitung:	Konvertierung von Excel 2013 nach PAULA-XML und ANNIS
Aufbereitungsart:	automatisch
Tool:	Pepper, Konvertierung des Formats Excel nach ANNIS und PAULA mithilfe der folgenden Importer: SpreadsheetImporter. Export ins ANNIS- und PAULA-Format mit dem ANNISExporter und dem PAULAExporter (https://korpling.german.hu-berlin.de/saltnpepper/pepper/download/stable/Pepper_2019.06.11.zip)
Format:	PAULA-XML, ANNIS
Segmentierung der Annotation:	basiert auf 'dipl'
Qualitätsprüfer:	korpling
Datum:	2019
Qualitätsprüfung:	NA
Editor:	Korpling

Typ:*Annotationlayer* – unclear

Annotationsart:	Annotationswerte werden je dipl-Einheit zugewiesen
Beschreibung:	Markierung, dass Zeichen eines Tokens oder ein Token bzw. ein Text nicht lesbar sind.

Typ: *Annotationvalue* – unclear

Wert:	Wertbeschreibung:
unclear	Nicht leserliche Zeichen oder Zeichenketten werden unabhängig von der Anzahl unleserlicher Zeichen, in der dipl-Ebene mit nur einem Unterstrich („_“) gekennzeichnet (siehe Transkription und Normalisierung). Wenn in der Transkription ein solcher Fall eintritt, so wird mit dem Tag „unclear“ in der gleichnamigen unclear-Ebene dies allgemein angezeigt. In die Ebenen **atLeast** und **atMost** wird angegeben, wieviele Zeichen „von“ „bis“ fehlen bzw. nicht gelesen werden können.

5.7. figure

Typ: *Layer* – figure

Name:	figure
Dokumentation:	Informationen zum gesamten Konvertierungsworkflow gibt es hier: https://github.com/langbank/conversion-pipeline.
Beschreibung:	Unter strukturellen Annotationen werden alle Annotationen zusammengefasst, die Aussagen über die graphische Strukturierung der Textgrundlage liefern. Dazu zählen Kennzeichnungen der Zeilen- und Seitenumbrüche, Seitenzahlen, unkenntliche/nicht lesbare Zeichen und Marginalien sowie Fußnoten. Strukturelle Annotationen werden auf Grundlage der diplomatischen Ebene vergeben (**dipl**-Ebene). So ist die Tokenisierung in der **dipl**-Ebene Vorlage für die weiteren, strukturellen Annotationen.

Typ: Preparationstep – figure

Schritt:	1
Aufbereitung:	Strukturelle Annotation
Aufbereitungsart:	manuell
Tool	Excel 2013
Format:	Excel 2013
Segmentierung der Annotation:	basiert auf ‘dipl’
Qualitätsprüfer:	korpling
Datum:	2019
Qualitätsprüfung:	NA
Editor:	Studierende der Humboldt-Universität zu Berlin und der Georg-August-Universität Göttingen, korpling

Schritt:	2
Aufbereitung:	Konvertierung von Excel 2013 nach PAULA-XML und ANNIS
Aufbereitungsart:	automatisch
Tool:	Pepper, Konvertierung des Formats Excel nach ANNIS und PAULA mithilfe der folgenden Importer: SpreadsheetImporter. Export ins

	ANNIS- und PAULA-Format mit dem ANNISExporter und dem PAULAExporter (https://korpling.german.hu-berlin.de/saltnpepper/pepper/download/stable/Pepper_2019.06.11.zip)
Format:	PAULA-XML, ANNIS
Segmentierung der Annotation:	basiert auf 'dipl'
Qualitätsprüfer:	korpling
Datum:	2019
Qualitätsprüfung:	NA
Editor:	Korpling

Typ: Annotationlayer – figure

Annotationsart:	Spannenannotation über einzelne oder mehrere dipl-Einheiten
Beschreibung:	Bisher wurden Abbildungen als Spanne über diejenigen dipl-Einheiten annotiert, mit denen sie auf einer Höhe sind, bzw. mit denen sie auf einer Seite stehen. Für die Dokumente ab Version 7 wurde eine präzisere Annotation vorgenommen: Markiert eine Graphik, die im Text eingebettet ist. Grundsätzlich wird unterschieden zwischen Abbildungen mit und ohne Bildbeschriftung. Besitzt eine Abbildung eine Beschriftung, werden alle dipl-Einheiten der Bildbeschriftung in einer Spanne als ‚figure' annotiert. Beispiele finden sich in der Tabelle „Abbildungen mit Beschriftung. Besitzt eine Abbildung keine Beschriftung, werden Teile des Fließtextes als „Anker" zur Annotation in der figure-Ebene verwendet. Diese sind in der Tabelle „Abbildungen ohne Beschriftung" definiert.

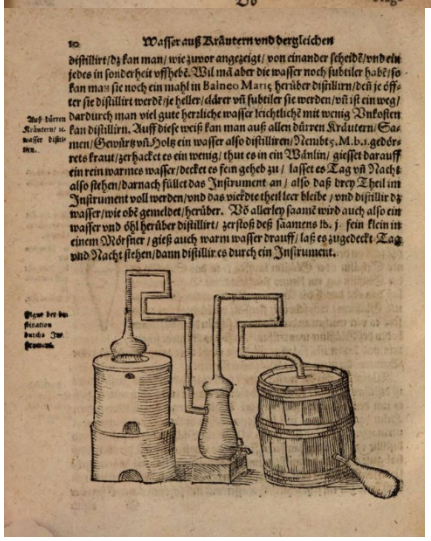
Abbildungen mit Beschriftung

Beschreibung	**Annotation**	**Screenshot**

Bild[1] mit Bildbeschriftung auf einer Extraseite	Die Bildbeschriftung wird wie der restliche Text auch transkribiert und bekommt eine eigene Seitenzahl zugewiesen. In ‚note' wird sie als *caption* annotiert und in ‚figure' als *figure*. In pb_n und in ‚figure_p' wird die entsprechende Seitenzahl angegeben. Sollte diese Extraseite einen Fließtext unterbrechen, d.h. nicht zwischen abgegrenzten Absätzen stehen, dann wird die Bildbeschriftung vor den entsprechenden Absatz gestellt und in pb_n ignoriert. In figure_p wird die Originalseitenzahl eingetragen. Bei dem hier aufgeführten Beispiel kann ein zweispaltiger Text angenommen werden (*Wal- wurtz* bekäme also in ‚column' den Wert *l* und *männ- lin* den Wert *r*). Auch Zeilenumbrüche werden annotiert.	
Zwei Bilder[2] mit Bildbeschriftung auf einer Extraseite	Die Bildbeschriftung wird wie der restliche Text auch transkribiert und bekommt eine eigene Seitenzahl zugewiesen. In ‚note' wird jede Beschriftung für sich als *caption* annotiert und in ‚figure' als *figure*. Es gibt also zwei *caption*- und zwei *figure*-Annotationen. In ‚figure_p' wird jeweils die entsprechende Seitenzahl angegeben. Bei dem hier aufgeführten Beispiel kann ein zweispaltiger Text angenommen werden (*Weiß Hornungsblüm* bekäme also in ‚column' den Wert *l* und *Geel Hornungsblüm* den Wert *r*).	

[1] Bild: Contrafayt Kreüterbuch / Otto Brunfels (Straßburg, 1532) Münchener DigitalisierungsZentrum http://mdz-nbn-resolving.de/urn:nbn:de:bvb:12-bsb00054201-5, Seite https://daten.digitale-sammlungen.de/bsb00054201/image_51

[2] Bild: Contrafayt Kreüterbuch / Otto Brunfels (Straßburg, 1532) Münchener DigitalisierungsZentrum http://mdz-nbn-resolving.de/urn:nbn:de:bvb:12-bsb00054201-5, Seite https://daten.digitale-sammlungen.de/bsb00054201/image_90

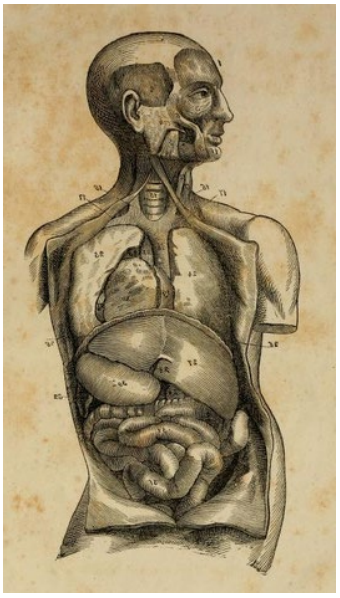
Mehr als zwei Bilder[3] mit Bildbeschriftung auf einer Extraseite	Bei mehreren Bildern wird von links nach rechts und von oben nach unten vorgegangen (links nach rechts geht vor). In diesem Beispiel werden die Figuren also in der folgenden Reihenfolge transkribiert und annotiert: - oben links - oben rechts - unten links - unten rechts werden jeweils in figure als „figure“ annotiert und erhalten in pb_n eine gemeinsame und in figure_p jeweils eine Seitenzahl zugewiesen	
Bild[4] mit Marginalie als Bildbeschriftung	Bezieht sich eine Marginalie offensichtlich auf eine Abbildung, die unter einem abgeschlossenen Absatz steht, wird diese Marginalie im Anschluss an den Absatz transkribiert und in ‚note‘ als *margin* sowie in ‚figure‘ als *figure* annotiert.	

[3] Bild: Phythologia/ Johann Joachim Becher (Ulm, 1662). Google Books https://books.google.de/books?id=-FlcAAAAcAAJ&pg=PA414&dq=phythologia&hl=de&sa=X&ei=4PiSVf7lAaH5ywObwaHABw&ved=0CCEQ6AEwAA#v=onepage&q&f=false

[4] Alchymistische Practic / Andreas Libavius (Frankfurt, 1603). Münchener DigitalisierungsZentrum http://mdz-nbn-resolving.de/urn:nbn:de:bvb:12-bsb10872546-0. Seite https://reader.digitale-sammlungen.de/de/fs1/object/display/bsb10872546_00016.html

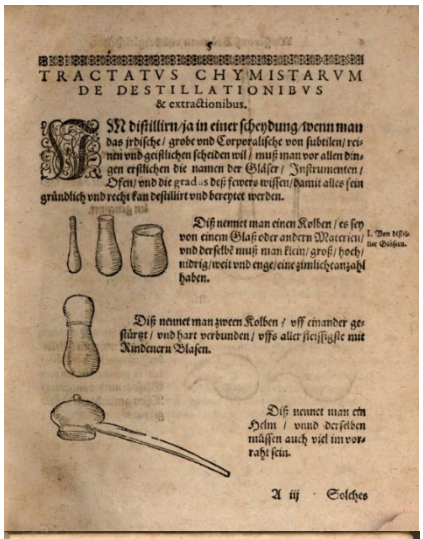
Zwei Bilder[5] mit Beschriftung zwischen Fließtext	Zuerst wird die Beschriftung des linken Bildes (oder des oberen, wenn übereinander), dann die des rechten Bildes vor dem entsprechenden Absatz transkribiert. Die Transkriptionsabschnitte werden in ‚note' jeweils als *caption* und in ‚figure' als *figure* annotiert. In ‚figure_p' wird die originale Seitenzahl angegeben. Bei dem konkreten Beispiel kann man außerdem eine Annotation von Spalten vornehmen.	

Abbildungen ohne Beschriftung

Beschreibung	**Annotation**	**Screenshot**
Bild(er) allein auf Seite[6]	Die letzte dipl-Einheit der vorangehenden sowie die erste dipl-Einheit der folgenden Seite werden als eine Figur annotiert. In ‚figure_p' wird die Seitenzahl angegeben. Sind auf einer Seite mehrere Abbildungen ohne Text werden diese als eine Figur annotiert.	

[5] Bild: Neuwkreuterbuch/ Jacobus Theodorus Tabernaemontanus (Frankfurt am Main, 1588). Collection Bnu en dépôt à l'Université de Strasbourg https://docnum.unistra.fr/digital/collection/coll13/id/7058

[6] Bild: Großes illustriertes Kräuterbuch / Ferdinand Müller (Ulm, 1860). Internet archive https://archive.org/details/dasgrosseillustr00muel/page/n6/mode/2up

Bilder[7] neben einspaltigem Text	Annotation von der ersten bis zur letzten dipl-Einheit des danebenstehenden Textes (wie bisher auch).	
Bild über Text[8]	Das Bild wird in ‚figure' für die erste dipl-Einheit des folgenden Textabschnittes annotiert (hier *Haſzelwurtz*).	
Bild unter Text	Das Bild wird in ‚figure' für die letzte dipl-Einheit des vorangehenden Textabschnittes annotiert.	

[7] Alchymistische Practic / Andreas Libavius (Frankfurt, 1603). Münchener DigitalisierungsZentrum http://mdz-nbn-resolving.de/urn:nbn:de:bvb:12-bsb10872546-0. Seite https://reader.digitale-sammlungen.de/de/fs1/object/display/bsb10872546_00011.html

[8] Bild: Contrafayt Kreüterbuch / Otto Brunfels (Straßburg, 1532) Münchener DigitalisierungsZentrum http://mdz-nbn-resolving.de/urn:nbn:de:bvb:12-bsb00054201-5, Seite https://daten.digitale-sammlungen.de/bsb00054201/image_47

Bild zwischen Text[9]	Das Bild wird in ‚figure' für die letzte dipl-Einheit des vorangehenden Textabschnittes und die erste dipl-Einheit des folgenden Textabschnittes annotiert (hier also für den Punkt hinter *Kuchenſchell* und *Kuchenſchell* im unteren Textabschnitt).	
Bild in zweispaltigem Text[10]	oben in einer Spalte: Annotation an der ersten nachfolgenden dipl-Einheit unten in einer Spalte: Annotation an der letzten vorangehenden dipl-Einheit innerhalb einer Spalte: Annotation an der letzten vorangehenden und ersten folgenden dipl-Einheit	

[9] Bild: Contrafayt Kreüterbuch / Otto Brunfels (Straßburg, 1532) Münchener DigitalisierungsZentrum http://mdz-nbn-resolving.de/urn:nbn:de:bvb:12-bsb00054201-5, Seite https://daten.digitale-sammlungen.de/bsb00054201/image_183

[10] Für drei folgende Bilder: Gart der Gesundheit / Johannes von Cuba (Ulm, 1487). Münchener DigitalisierungsZentrum http://mdz-nbn-resolving.de/urn:nbn:de:bvb:12-bsb00048197-0. Seite https://daten.digitale-sammlungen.de/bsb00048197/image_6, https://daten.digitale-sammlungen.de/bsb00048197/image_8, https://daten.digitale-sammlungen.de/bsb00048197/image_20

die feüchtblatœn darauff gelegt
oß auf beülle hind den oœ/daenoe
so sol man dz mit ainœ schrepffli-
ten blütend machē vñ darauff dz
bulfœ legen¶ Item beyfuß ist fœ
best opfung dr gelider auf tün als
milcz vnd lebern die do von ainer
kaltē materiē komet also gemüczt
Nim beyfuß vnd hirßzügen yeg-
liche ain hand vol vñ wenig wer
mūt in wein gesotē vñ mit zuckr-
süß gemachet.¶ Diser trancк ist
auch gůt wid die gelsucht/ so mā
darzu ist vermögen wenig tausēt
guldin genāt centaurea· vñ was
nach dem getrancк überbleibt wō
kreütern sol man warm legen auf
die milcz/Das selb ist auch gut so
die milcz aufwendig geschwollē
ist.¶ Item beyfuß vñ walle genāt
tazus barbatus gesoten mit wein
vnnd also warm darauf gesessen
ist gut dem drfdarm außgat.

Stabwurtz
Das ander Capitel

Brotunum· Das wort ist
in grecum vnd in latinitz
also genant· Hesum ist es
genant in arabischem¶ Dr mai
stœr Sœapio in dem büch aggre-
gatoris in dē capitel Hesam id est
Abrotunum spricht· das diß kraut
sey zweyer hand/ ains ist fœülich
das and mānlich/ vnd sind baide
ainander geleich an dr gestalt vñ
an dr tugendt· allain die fœülich
hat weiß blümē/ vnd die mānlich
gelb blümen vnd geleich dem saf-
fran.¶ Dr maistœr Plinius in sei
nem xxj· büch spricht· dz stab-
wurtz hab ainen gůten geruch.
¶ Diß kraut hat fast kleine blet-
tœr vnd vil stengel/ vnd auch gar
klain blümē vnd ain klainē samē
¶ Dr maistœr Auicēna in seinem
andn büch spricht· dz stabwurtz
sey haiß an dem erſten grade vnd
trucken an dem andern¶ Dr wir
dig maistœr Auicenna in seinē an-
dren büch in dē capitel Abrotunū
spricht· das stabwurtz gebrennet
vnd gebulfert vnd das gemischet
mit öl wō rätich gemacht vñ das
geschmieret auf die kal stat macht
har wachsen¶ Item stabwurtz
macht auch wachsen den bart ō
langesam herfür komet do mit be-
strichē/ mit namē nym stabwurtz
safft mit dyll öle vermengt/ dañ
den bart oder die stat geschmieret
do man har begert¶ Stabwurtz

gestosse/ vnd mit dyptam wurtz
gebulfert/ mit hōnig vermengt in
ainer salben weise/ vñ auf wundē
gelegt/ ist beulen vnd dornen auß
der wunden ziehen.

Osterlucien
Das vi Capitel

Aristologia longa· ¶ Die
wirdige maistœr Auicēna
Galien⁹ Diascorides pla
tearius vnd Plinius sprechē· das
die lang holwurtz seye warm an
dem dritten grad/ vnnd trucken an
dē andn grad· Die lang holwurtz
hat ain langen stil vnd bletœr dar
an die geleichē den bletœrn an der
haselwurtz/ allain das osterlucie
oß holwurtz bletœr weichœ sind
an dem griff· Die wurtzel ist lang
geleich den peterſiligten wurtzeln·
¶ Plinius spricht· das baide oster
lucien die lang vñ die rund gebul
fœt vnd darunder gemischt mitrē
yeglichs geleich vil/ vnd disen ge-
mischt ain quintin mit warmē wei
ninget die mūttœr genant matrix
von irer vnfläßigkait/ vnnd treibt
auß die tote geburt¶ Für das po
dogram nym wegbraite eybisch-
wurtzeln vnd lang holwurtzeln
yeglichs geleich vil vñ misch dar-
under hōnig vnd leg das auff den
gebresten es senftiget vñ mindert
den wetagen¶ Die osterlucie ist
fast gůt den faulē wunden des bul
fœs darein gestreut.¶ Welches
pfärd gewundet wirt vō vil reitē
oder tragen der strewe des bulfœs
von disē osterlucie in die wundē/
vnd schlahe darauff pfärdes mist
es hailet tzehand.¶ Platearius-
nym osterlucie vnd aloepaticum
yeglichs geleich vil vnd misch dar
under rosen hōnig vnd mach dar
auß ain pflastœr· disen pflastœr ist
gůt genützet allen alten schäden/
als dañ ist der krebß fistel ō wolff
wie die sein mügen an den bainen
vnd anderſwo· hailet es senftig-
lich darauf gelegt· vñ die schädē
oder gebresten sollen vorhin gesu
beret werden mit weinstain öl oō
mit wosser dariñ mirra gesotē ist.
¶ Diß also genützet benymet die
ausserzigkait.¶ Diascorides oster-
lucie gebulfœt vnd gemischet mit
hōnigwasser vnd das getrunckē
benimt ofma/ das ist das ticchen/
vnd rainiget die brust.¶ Dr ain
pfeil in seinē leib hęt/ō neme baid

Bild[11] mittig zwischen zweispaltigem Text; **Lesefluss**: links rechts, links rechts	Das Bild wird für die letzte dipl-Einheit in der rechten Spalte oben bis zur ersten dipl-Einheit in der linken Spalte unten annotiert.	

[11] Bild: Wund-Artzney / Friedrich Greiff, Guilelmus Fabricius Hildanus (Frankfurt, 1652). Münchener DigitalisierungsZentrum http://mdz-nbn-resolving.de/urn:nbn:de:bvb:12-bsb10726741-5. https://reader.digitale-sammlungen.de/de/fs1/object/display/bsb10726741_00049.html

Bild[12] mittig zwischen zweispaltigem Text; **Lesefluss**: links, rechts	Das Bild wird für die letzte dipl-Einheit in der linken Spalte oben und die erste dipl-Einheit in der linken Spalte unten annotiert.	

[12] Bild: Wund-Artzney / Friedrich Greiff, Guilelmus Fabricius Hildanus (Frankfurt, 1652). Münchener DigitalisierungsZentrum http://mdz-nbn-resolving.de/urn:nbn:de:bvb:12-bsb10726741-5. https://reader.digitale-sammlungen.de/de/fs1/object/display/bsb10726741_00049.html

Bild[13] mittig über zweispaltigem Text/ Bild unten in der linken Spalte	Das obere Bild wird an der ersten dipl-Einheit der linken Spalte annotiert, das untere Bild für die nebenstehenden dipl-Einheiten.	
Bild von Text umgeben	Das Bild wird für den gesamten Text auf dieser Seite annotiert.	T1 T2 T3 T 4 **B** T5 T6 T7 T8

Typ: Annotationvalue – figure

Wert:	Wertbeschreibung:
figure	Allgemeine Kennzeichnung einer Grafik im Text.
table	Allgemeine Kennzeichnung einer Tabelle im Text

5.8. figure_p

Typ: *Layer* – figure_p

Name:	figure_p
Dokumentation:	Informationen zum gesamten Konvertierungsworkflow gibt es hier: https://github.com/langbank/conversion-pipeline.

[13] Bild: Wund-Artzney / Friedrich Greiff, Guilelmus Fabricius Hildanus (Frankfurt, 1652). Münchener DigitalisierungsZentrum http://mdz-nbn-resolving.de/urn:nbn:de:bvb:12-bsb10726741-5. https://reader.digitale-sammlungen.de/de/fs1/object/display/bsb10726741_00053.html

Beschreibung:	Unter strukturellen Annotationen werden alle Annotationen zusammengefasst, die Aussagen über die graphische Strukturierung der Textgrundlage liefern. Dazu zählen Kennzeichnungen der Zeilen- und Seitenumbrüche, Seitenzahlen, unkenntliche/nicht lesbare Zeichen und Marginalien sowie Fußnoten. Strukturelle Annotationen werden auf Grundlage der diplomatischen Ebene vergeben (**dipl**-Ebene). So ist die Tokenisierung in der **dipl**-Ebene Vorlage für die weiteren, strukturellen Annotationen.

Typ: Preparationstep – figure_p

Schritt:	1
Aufbereitung:	Strukturelle Annotation
Aufbereitungsart:	manuell
Tool	Excel 2013
Format:	Excel 2013
Segmentierung der Annotation:	basiert auf ‘dipl’
Qualitätsprüfer:	korpling
Datum:	2019
Qualitätsprüfung:	NA
Editor:	Studierende der Humboldt-Universität zu Berlin und der Georg-August-Universität Göttingen, korpling

Schritt:	2
Aufbereitung:	Konvertierung von Excel 2013 nach PAULA-XML und ANNIS
Aufbereitungsart:	automatisch
Tool:	Pepper, Konvertierung des Formats Excel nach ANNIS und PAULA mithilfe der folgenden Importer: SpreadsheetImporter. Export ins ANNIS- und PAULA-Format mit dem ANNISExporter und dem PAULAExporter (https://korpling.german.hu-berlin.de/saltnpepper/pepper/download/stable/Pepper_2019.06.11.zip)
Format:	PAULA-XML, ANNIS
Segmentierung der Annotation:	basiert auf ‘dipl’
Qualitätsprüfer:	korpling
Datum:	2019

Qualitätsprüfung:	NA
Editor:	Korpling

Typ: *Annotationlayer* – figure_p

Annotationsart:	Spannenannotation über mehrere dipl-Einheiten
Beschreibung:	Angabe der Seite, auf der eine Abbildung im Original zu sehen ist.

5.9. script

Typ: *Layer* – 5.12. script

Name:	script
Dokumentation:	Informationen zum gesamten Konvertierungsworkflow gibt es hier: https://github.com/langbank/conversion-pipeline.
Beschreibung:	Unter strukturellen Annotationen werden alle Annotationen zusammengefasst, die Aussagen über die grafische Strukturierung der Textgrundlage liefern. Dazu zählen Kennzeichnungen der Zeilen- und Seitenumbrüche, Seitenzahlen, unkenntliche/nicht lesbare Zeichen und Marginalien sowie Fußnoten. Strukturelle Annotationen werden auf Grundlage der diplomatischen Ebene vergeben (**dipl**-Ebene). So ist die Tokenisierung in der **dipl**-Ebene Vorlage für die weiteren, strukturellen Annotationen.

Typ:*Preparationstep* – script

Schritt:	1
Aufbereitung:	Strukturelle Annotation
Aufbereitungsart:	manuell
Tool	Excel 2013
Format:	Excel 2013
Segmentierung der Annotation:	basiert auf ‘dipl’
Qualitätsprüfer:	korpling
Datum	2019

Qualitätsprüfung:	NA
Editor:	Studierende der Humboldt-Universität zu Berlin und der Georg-August-Universität Göttingen, korpling
Schritt:	2
Aufbereitung:	Konvertierung von Excel 2013 nach PAULA-XML und ANNIS
Aufbereitungsart:	automatisch
Tool:	Pepper, Konvertierung des Formats Excel nach ANNIS und PAULA mithilfe der folgenden Importer: SpreadsheetImporter. Export ins ANNIS- und PAULA-Format mit dem ANNISExporter und dem PAULAExporter (https://korpling.german.hu-berlin.de/saltnpepper/pepper/download/stable/Pepper_2019.06.11.zip)
Format:	PAULA-XML, ANNIS
Segmentierung der Annotation:	basiert auf 'dipl'
Qualitätsprüfer:	korpling
Datum:	2019
Qualitätsprüfung:	NA
Editor:	Korpling

Typ:*Annotationlayer* – script

Annotationsart:	Spannenannotation für einzelne dipl-Einheiten
Beschreibung:	Weist jedem Token eine Schriftart zu. In manchen Texten kommt es vor, dass ein und dasselbe Token zum Teil in einem und zum Teil in einem anderen Font gedruckt wurde. Dies betrifft v.a. Wörter mit deutscher Flexionsendung und fremdsprachiger Wurzel. Für diese Token wurde innerhalb von „script" der Wert „mixed" annotiert.

Typ:*Annotationvalue* – script

Wert:	Wertbeschreibung:
roman	Text, der in Antiqua gedruckt ist.

blackletter	Text, der in Fraktur gedruckt ist.
mixed	Text, der in verschiedenen Schriftsätzen gedruckt ist.

5.10. hi_rend

Typ: *Layer* – hi_rend

Name:	hi_rend
Dokumentation:	Informationen zum gesamten Konvertierungsworkflow gibt es hier: https://github.com/langbank/conversion-pipeline.
Beschreibung:	Unter strukturellen Annotationen werden alle Annotationen zusammengefasst, die Aussagen über die graphische Strukturierung der Textgrundlage liefern. Dazu zählen Kennzeichnungen der Zeilen- und Seitenumbrüche, Seitenzahlen, unkenntliche/nicht lesbare Zeichen und Marginalien sowie Fußnoten. Strukturelle Annotationen werden auf Grundlage der diplomatischen Ebene vergeben (**dipl**-Ebene). So ist die Tokenisierung in der **dipl**-Ebene Vorlage für die weiteren, strukturellen Annotationen.

Typ:*Preparationstep* – hi_rend

Schritt:	1
Aufbereitung:	Strukturelle Annotation
Aufbereitungsart:	manuell
Tool	Excel 2013
Format:	Excel 2013
Segmentierung der Annotation:	basiert auf ‘dipl’
Qualitätsprüfer:	korpling
Datum	2018
Qualitätsprüfung:	Manuelle Konsistenzprüfung: Korrektur der false positives
Editor:	Studierende der Humboldt-Universität zu Berlin und der Georg-August-Universität Göttingen
Schritt:	2

Aufbereitung:	Konvertierung von Excel 2013 nach PAULA-XML und ANNIS
Aufbereitungsart:	automatisch
Tool:	Pepper, Konvertierung des Formats Excel nach ANNIS und PAULA mithilfe der folgenden Importer: SpreadsheetImporter. Export ins ANNIS- und PAULA-Format mit dem ANNISExporter und dem PAULAExporter (https://korpling.german.hu-berlin.de/saltnpepper/pepper/download/stable/Pepper_2019.06.11.zip)
Format:	PAULA-XML, ANNIS
Segmentierung der Annotation:	basiert auf 'dipl'
Qualitätsprüfer:	korpling
Datum:	2019
Qualitätsprüfung:	NA
Editor:	Korpling

Typ:*Annotationlayer* – hi_rend

Annotationsart:	Spannenannotation über einzelne dipl-Einheiten
Beschreibung:	Beschreibung der Darstellung des hervorgehobenen Bereichs, z. B. graphische Hervorhebungen, Farben etc.

Typ:*Annotationvalue* – hi_rend

Wert:	Wertbeschreibung:
italics	Text, der kursiv gedruckt ist.
bold	Text, der fett gedruckt ist.
underlined	Text, der unterstrichen gedruckt ist.
red	Text, der rot gedruck ist.
iniCap	Ein Text, der mit einem zusätzlichen Großbuchstaben gedruckt ist.
letter-spacing:1em	Ein größerer Abstand zwischen Buchstaben, Sperrdruck.

5.11. head

Typ: *Layer* – head

Name:	head
Dokumentation:	Informationen zum gesamten Konvertierungsworkflow gibt es hier: https://github.com/langbank/conversion-pipeline.
Beschreibung:	Unter strukturellen Annotationen werden alle Annotationen zusammengefasst, die Aussagen über die graphische Strukturierung der Textgrundlage liefern. Dazu zählen Kennzeichnungen der Zeilen- und Seitenumbrüche, Seitenzahlen, unkenntliche/nicht lesbare Zeichen und Marginalien sowie Fußnoten. Strukturelle Annotationen werden auf Grundlage der diplomatischen Ebene vergeben (**dipl**-Ebene). So ist die Tokenisierung in der **dipl**-Ebene Vorlage für die weiteren, strukturellen Annotationen.

Typ: Preparationstep – head

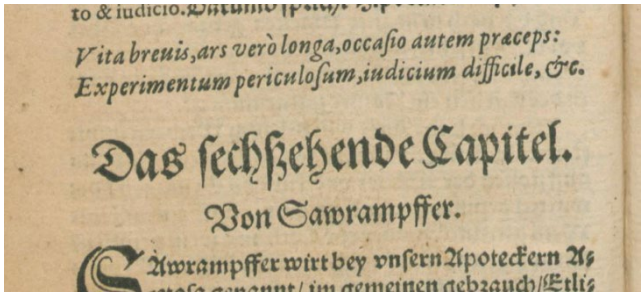
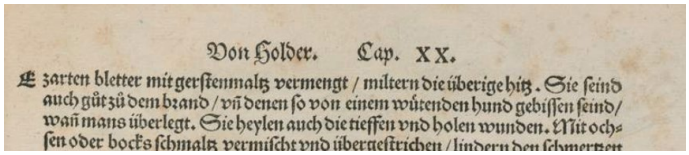
Schritt:	1
Aufbereitung:	Strukturelle Annotation
Aufbereitungsart:	manuell
Tool	Excel 2013
Format:	Excel 2013
Segmentierung der Annotation:	basiert auf ‘dipl’
Qualitätsprüfer:	LAUDATIO
Datum:	2019
Qualitätsprüfung:	NA
Editor:	Studierende der Humboldt-Universität zu Berlin und der Georg-August-Universität Göttingen

Schritt:	2
Aufbereitung:	Konvertierung von Excel 2013 nach PAULA-XML und ANNIS
Aufbereitungsart:	automatisch
Tool:	Pepper, Konvertierung des Formats Excel nach ANNIS und

	PAULA mithilfe der folgenden Importer: SpreadsheetImporter. Export ins ANNIS- und PAULA-Format mit dem ANNISExporter und dem PAULAExporter (https://korpling.german.hu-berlin.de/saltnpepper/pepper/download/stable/Pepper_2019.06.11.zip)
Format:	PAULA-XML, ANNIS
Segmentierung der Annotation:	basiert auf 'dipl'
Qualitätsprüfer:	korpling
Datum:	2019
Qualitätsprüfung:	NA
Editor:	Korpling

Typ: Annotationlayer – head

Annotationsart:	Spannenannotation über mehrere dipl-Einheiten
Beschreibung:	Markierung einer Überschrift, wenn sich diese vom Fließtext absetzt. Fortlaufende Kapitelüberschriften werden dabei nicht transkribiert (siehe Abbildungen).

Annotation	**Screenshot**
Zwei Überschriften weden jeweils in einer Spanne annotiert.[14]	
Eine fortlaufende Kapitelüberschrift wird nicht transkribiert.[15]	

Typ: Annotationvalue – head

[14] Bild: Experimenta von zwentzig Pestilenz Wurtzelen unnd Kreutter/ Tarquinius Schellenberg (Strassburg, 1550) Sammlung Vester (DFG) http://digital.ub.uni-duesseldorf.de/vester/content/pageview/1461389

[15] Bild: New Kreüterbuch (Kapitel 19-25) / Leonhart Fuchs (Basel, 1543). Münchener DigitalisierungsZentrum http://mdz-nbn-resolving.de/urn:nbn:de:bvb:12-bsb00017437-8. Seite https://daten.digitale-sammlungen.de/bsb00017437/image_94

Wert:	Wertbeschreibung:
head	Allgemeine Markierung einer Überschrift.

5.12. note

Typ: *Layer* – note

Name:	note
Dokumentation:	Informationen zum gesamten Konvertierungsworkflow gibt es hier: https://github.com/langbank/conversion-pipeline.
Beschreibung:	Unter strukturellen Annotationen werden alle Annotationen zusammengefasst, die Aussagen über die graphische Strukturierung der Textgrundlage liefern. Dazu zählen Kennzeichnungen der Zeilen- und Seitenumbrüche, Seitenzahlen, unkenntliche/nicht lesbare Zeichen und Marginalien sowie Fußnoten. Strukturelle Annotationen werden auf Grundlage der diplomatischen Ebene vergeben (**dipl**-Ebene). So ist die Tokenisierung in der **dipl**-Ebene Vorlage für die weiteren, strukturellen Annotationen.

Typ: Preparationstep – note

Schritt:	1
Aufbereitung:	Strukturelle Annotation
Aufbereitungsart:	manuell
Tool	Excel 2013
Format:	Excel 2013
Segmentierung der Annotation:	basiert auf ‘dipl’
Qualitätsprüfer:	korpling
Datum:	2019
Qualitätsprüfung:	NA
Editor:	Studierende der Humboldt-Universität zu Berlin und der Georg-August-Universität Göttingen

Schritt:	2
Aufbereitung:	Konvertierung von Excel 2013 nach PAULA-XML und ANNIS
Aufbereitungsart:	automatisch
Tool:	Pepper, Konvertierung des Formats Excel nach ANNIS und PAULA mithilfe der folgenden Importer: SpreadsheetImporter. Export ins ANNIS- und PAULA-Format mit dem ANNISExporter und dem PAULAExporter (https://korpling.german.hu-berlin.de/saltnpepper/pepper/download/stable/Pepper_2019.06.11.zip)
Format:	PAULA-XML, ANNIS
Segmentierung der Annotation:	basiert auf 'dipl'
Qualitätsprüfer:	korpling
Datum:	2019
Qualitätsprüfung:	NA
Editor:	Korpling

Typ: Annotationlayer – note

Annotationsart:	Spannenannotation über mehrere dipl-Einheiten
Beschreibung:	In einer Spannenannotation wird die gesamte Fußnote mit dem Tag „note" markiert. Der Text der Fußnoten wird immer am Ende des Absatzes, in dem sie eingepflegt sind, unabhängig davon, ob sich so die Fußnoten sammeln oder erst auf der nachfolgenden Seite realisiert werden müssten, in die dipl-Ebene/Transkription aufgenommen.

Typ: Annotationvalue – note

Wert:	Wertbeschreibung:
note	In einer Spannenannotation wird die gesamte Fußnote mit dem Tag „note" markiert. Der Text der Fußnoten wird immer am Ende des Absatzes, in dem sie eingepflegt sind, unabhängig davon, ob sich so die Fußnoten sammeln oder erst auf der nachfolgenden Seite realisiert werden müssten, in die dipl-Ebene/Transkription aufgenommen.
margin	In einer Spannenannotation wird die gesamte Marginalie mit dem Tag „margin" markiert. Der Text der Marginalien wird immer am Anfang des jeweiligen Absatzes, in dem/ neben dem sie realisiert sind, unabhängig davon, ob so mehrere

	Marginalien hintereinander oder/und auf der vorherigen Seite realisiert werden müssen, in die dipl-Ebene/Transkription aufgenommen.
caption	In einer Spannenannotation wird die gesamte Bildbeschriftung mit dem Tag „caption“ markiert. Der Text der Bildbeschriftungen wird immer am Anfang des jeweiligen Absatzes, in dem sie realisiert sind, unabhängig davon, ob so mehrere Bildbeschriftungen hintereinander oder/und auf der vorherigen Seite realisiert werden müssen, in die dipl-Ebene/Transkription aufgenommen.
end	Eine Endnote.

5.13. quote

Typ: *Layer* – quote

Name:	quote
Dokumentation:	Informationen zum gesamten Konvertierungsworkflow gibt es hier: https://github.com/langbank/conversion-pipeline.
Beschreibung:	Unter strukturellen Annotationen werden alle Annotationen zusammengefasst, die Aussagen über die graphische Strukturierung der Textgrundlage liefern. Dazu zählen Kennzeichnungen der Zeilen- und Seitenumbrüche, Seitenzahlen, unkenntliche/nicht lesbare Zeichen und Marginalien sowie Fußnoten. Strukturelle Annotationen werden auf Grundlage der diplomatischen Ebene vergeben (**dipl**-Ebene). So ist die Tokenisierung in der **dipl**-Ebene Vorlage für die weiteren, strukturellen Annotationen.

Typ: Preparationstep – quote

Schritt:	1
Aufbereitung:	Strukturelle Annotation
Aufbereitungsart:	manuell
Tool	Excel 2010
Format:	Excel 2010
Segmentierung der Annotation:	basiert auf ‘dipl’
Qualitätsprüfer:	LAUDATIO

Datum:	2019
Qualitätsprüfung:	NA
Editor:	Studierende der Humboldt-Universität zu Berlin und der Georg-August-Universität Göttingen

Schritt:	2
Aufbereitung:	Konvertierung von Excel 2013 nach PAULA-XML und ANNIS
Aufbereitungsart:	automatisch
Tool:	Pepper, Konvertierung des Formats Excel nach ANNIS und PAULA mithilfe der folgenden Importer: SpreadsheetImporter. Export ins ANNIS- und PAULA-Format mit dem ANNISExporter und dem PAULAExporter https://korpling.german.hu-berlin.de/saltnpepper/pepper/download/stable/Pepper_2019.06.11.zip)
Format:	PAULA-XML, ANNIS
Segmentierung der Annotation:	basiert auf 'dipl'
Qualitätsprüfer:	korpling
Datum:	2019
Qualitätsprüfung:	NA
Editor:	Korpling

Typ: Annotationlayer – quote

Annotationsart:	Spannenannotation für einzelne dipl-Einheiten
Beschreibung:	Ordnet Fließtext als Zitat oder Nicht-Zitat ein.

Typ: Annotationvalue – quote

Wert:	Wertbeschreibung:
yes	Ein Zitat im Text.
no	Kein Zitat (Default-Wert).

6. Annotationsebenen – Inhaltliche Annotation

6.1. persname

Typ: *Layer* – persname

Name:	persname
Dokumentation:	Es werden alle Personennamen annotiert, auf die ein Autor referiert. Es wird der jeweilige Personenname im Nominativ angegeben. Eine Liste der vereinheitlichten Namen findet sich unten. Informationen zum gesamten Konvertierungsworkflow gibt es hier: https://github.com/langbank/conversion-pipeline.
Beschreibung:	Unter linguistischen Annotationen werden u.a. pos-Tagging, Lemmatisierung und Markierung von fremdsprachlichem Material zusammengefasst.

Typ: Preparationstep – persname

Schritt:	1
Aufbereitung:	Inhaltliche Annotation
Aufbereitungsart:	manuell
Tool	Excel 2013
Format:	Excel 2013
Segmentierung der Annotation:	basiert auf ‘dipl’
Qualitätsprüfer:	korpling
Datum	2018
Qualitätsprüfung:	manuelle Korrektur
Editor:	Studierende der Humboldt-Universität zu Berlin und der Georg-August-Universität Göttingen, korpling

Schritt:	2
Aufbereitung:	Konvertierung von Excel 2013 nach PAULA-XML und ANNIS
Aufbereitungsart:	automatisch
Tool:	Pepper, Konvertierung des Formats Excel nach ANNIS und PAULA mithilfe der folgenden Importer: SpreadsheetImporter. Export ins ANNIS- und PAULA-Format mit dem ANNISExporter und dem PAULAExporter (https://korpling.german.hu-berlin.de/saltnpepper/pepper/download/stable/Pepper_2019.06.11.zip)

Format:	PAULA-XML, ANNIS
Segmentierung der Annotation:	basiert auf ‘dipl’
Qualitätsprüfer:	korpling
Datum:	2019
Qualitätsprüfung:	NA
Editor:	Korpling

Typ: Annotationlayer – persname

Annotationsart:	Spannenannotation über eine oder mehrere norm-Einheiten
Beschreibung:	Es werden alle Personennamen annotiert, auf die ein Autor referiert.

Typ: Annotationvalue – persname

Wert:	Wertbeschreibung:
Zeichenkette	Diese Ebene dient der eindeutigen und schnellen Verknüpfung der im Korpus auftretenden Namen mit den dazugehörigen historischen Personen. Ist eine Person nicht eindeutig identifizierbar wird der in der "norm" gegebene Wert im Nominativ Singular in die Ebene "persname" übertragen. Die Schreibung der Namen richtet sich nach dieser Liste (diese enthält offizielle Abkürzungen und die entsprechende Ausschreibung vieler BotanikerInnen) und wenn dort nicht aufgeführt nach ihrer Schreibung in der deutschsprachigen Wikipedia. Anstelle von Leerzeichen werden Unterstriche gesetzt. Im Folgenden sind alle bisher in RIDGES vorkommenden **eindeutigen** Personennamen aufgeführt. Diese werden weiterhin so übernommen. Hinweis: Viele Autoren haben eine Art Künstlername und es gibt auch viele verschiedene Schreibweisen. In dieser Ebene muss darauf geachtet werden, dass ein und dieselbe Person immer mit dem gleichen Namen annotiert wird.

A Achilles, Adriaan_van_Royen, Aetios_von_Amida, Agrippa_von_Nettesheim, Aimé_Jacques_Alexandre_Bonpland, Albrecht_von_Haller, Alexander_von_Humboldt, Alexander_von_Tralleis, Alexander_Wilhelm_Martini, Alexandre_Henri_Gabriel_de_Cassini, al_Farabi, Almanzor, André_Michaux, Andreas_Cleyer, Anna_von_Böhmen_und_Ungarn, Antoine-Laurent_de_Jussieu, Antoine_Gouan, Anton_Rehmann, Antonio_Musa_Brassavola, Antonius_Musa, Apollonius, Apuleius, Archigenes, Aristomachos, Aristoteles, Arnald_von_Villanova, Artemis, Asa_Grey,

Asklepiades_von_Bithynien, Attalos_I., August_Quirinus_Rivinus,
Augustin_Pyramus_De_Candolle, Averroes, Avicenna, **B** Balthasar_Ehrhart,
Bartholomaeus_Anglicus, Bartolomeo_Platina, **C** Caelius_Apicius,
Carl_Heinrich_Schultz_Bipontinus, Carl_Peter_Thunberg, Carl_Sigismund_Kunth,
Carl_von_Linné, Carrichter, Caspar_Bauhin, Caspar_Hoffmann, Charles_de_l'Écluse,
Christian_Friedrich_Lessing, Christian_Gottlieb_Ludwig, Christoph_Wilhelm_Hufeland,
Chrysippos_von_Soloi, Conrad_Gessner, Conrad_Roßbach, Constantinus, **D**
David_Heinrich_Hoppe, Dieuches, Diodoros, Dirslaus_von_Schwenkfeld,
Dominique_Villars, **E** Eberhard_Gmelin, Eduard_August_von_Regel, Erasistratos,
Eucharius_Rösslin_der_Jüngere, **F** Fabio_Colonna, Ferdinand_Jacob_Heinrich_Sir_Müller,
Filippo_Beroaldo, Flavius_Josephus, François_Fulgis_Chevallier, Franz_Xaver_von_Wulfen,
Franziskus_Joel, Friedrich_II., Friedrich_III., **G** Galenos, Gaza, Genthios,
Georg_Franz_Hoffmann, George_Bentham, Giovanni_da_Vigo, Gregor_Horst,
Guglielmo_da_Varignana, Guillaume_Rondelet, Gustav_Karl_Wilhelm_Karsten,
Guy_de_La_Brosse, **H** Heinrich_Bernhard_Rupp, Heinrich_Kramer, Heinrich_Petreus,
Henri_Ernst_Baillon, Henry_Charles_Andrews, Herakleides_Pontikos, Herkules,
Herman_Boerhaave, Hermann_Hager, Hermann_von_Neuenahr_der_Ältere,
Hermolaus_Barbarus, Hesiod, Hieronymus_Bock, Hipólito_Ruiz_López,
Hippokrates_von_Kos, Homer, Hugh_Algernon_Weddell, **I** Isaac_Newton, Isaac_Rand,
Isaak_ben_Salomon_Israeli, Isidor_von_Sevilla, **J** Jacques_Barrelier, Jacques_Peschier,
Jean-Baptiste_Miroudot_du_Bourg, Jean_Louis_Lassaigne, Jean_Ruel, Joachim_Camerarius,
Jesus_von_Nazaret, Joachim_Camerarius_der_Jüngere, João_de_Loureiro, Johann_Bauhin,
Johann_Esaias_Silberschlag, Johann_Georg_Volkamer, Johann_Gottlieb_Gleditsch,
Johann_Heinrich_von_Heucher, Johann_Jacob_Dillen, Johann_Jacob_Reinhard,
Johann_Jakob_Scheuchzer, Johann_Schröder, Johann_Sigismund_Elsholtz,
Johann_Wonnecke_von_Kaub, Johannes_Andrä_von_Königgrätz,
Johannes_Bodaeus_van_Stapel, Johannes_Jacobus_Manlius_de_Bosco, Johannes_Manardus,
Johannes_Mesuë_der_Ältere, Johannes_Theodor_Schmalhausen, Johannes_von_Damaskus,
John_Gerard, John_Hutton_Balfour, John_Kirk, John_Parkinson,
José_Antonio_Pavón_y_Jiménez, Joseph_Pitton_de_Tournefort,
Juan_José_Martínez_de_Lexarza, Juba_I., Justina, **K** Karl_August_von_Bergen,
Karl_Gottfried_Hagen, Antimachos_von_Kolophon, König_Salomo,
Königin_von_Halikarnassos, **L** Leonhart_Fuchs, Lorenz_Heister,
Lucius_Iunius_Moderatus_Columella, Lysimachos, **M** Marcellus_Vergilius,
Marquard_von_Hattstein, Maimonides, Marcellus_Empiricus, , Michael_Isigrin,
Marcus_Porcius_Cato_der_Ältere, Marschall_von_Bieberstein, Marsilio_Ficino,
Martin_Schmuck, Matteo_Syluatico, Matthaeus_Platearius, Mausolos_II, Maximilian_I.,
Michael_Ettmüller, Michael_Meyenburg, Michelangelo_Tilli, Mithridates, **N**
Nathaniel_Wallich, Nero, Niccolò_Leoniceno, Nicolas_Lémery, Nicolaus_Joseph_Jacquin,
Nikandros_aus_Kolophon, Nikolai_Stepanowitsch_Turtschaninow, **O** Oreibasios, Orpheus,
Otto_Brunfels, Otto_Heinrich, **P** Pablo_de_La_Llave, Pandecta, Paulos_von_Aigina, Paulus,
Paulus_der_Arzt, Paracelsus, Paulus_von_Aegineta, Paulus_von_Tarsus,
Pedanios_Dioskurides, Pehr_Kalm, Philip_Miller, Pierre_Séguier , Pietro_Andrea_Mattioli,
Platon, Plinius_der_Ältere, Pythagoras, **R** Raius, Ramon_Llull, Rhazes, Rembert_Dodoens,
Robert_Brown, Robert_Morison, Rudolf_Goclenius, Rutilius_Taurus_Aemilianus_Palladius,
S Samuel_Ledel, Sebastian_Kneipp, Serapion_von_Alexandria, Serenus,
Simon_Pauli_der_Jüngere, Simon_Petrus, Simon_von_Genua, Sosimenes,

Stanislaus_I._Leszczyński, Strabon, **T** Tabernaemontanus, Teukros_(König_von_Troas), Thales,Themison_von_Laodikeia, Theophilus_Protospatharius, Theophrastos_von_Eresos, Thomas_Nuttall, **W** Wilhelm_Daniel_Joseph_Koch, Wilhelm_Ernst_Scheffer, William_Henry_Harvey, William_Jackson_Hooker, William_Watson, **X** Xenokrates, Xenophon, **Z** Zenon

6.2. title

Typ: *Layer* – title

Name:	title
Dokumentation:	Informationen zum gesamten Konvertierungsworkflow gibt es hier: https://github.com/langbank/conversion-pipeline.
Beschreibung:	Unter linguistischen Annotationen werden u.a. pos-Tagging, Lemmatisierung und Markierung von fremdsprachlichem Material zusammengefasst.

Typ: Preparationstep – title

Schritt:	1
Aufbereitung:	Inhaltliche Annotation
Aufbereitungsart:	manuell
Tool	Excel 2013
Format:	Excel 2013
Segmentierung der Annotation:	basiert auf ‘dipl’
Qualitätsprüfer:	NA
Datum	2019
Qualitätsprüfung:	NA
Editor:	Studierende der Humboldt-Universität zu Berlin und der Georg-August-Universität Göttingen; korpling

Schritt:	2
Aufbereitung:	Konvertierung von Excel 2013 nach PAULA-XML und ANNIS
Aufbereitungsart:	automatisch
Tool:	Pepper, Konvertierung des Formats Excel nach ANNIS und PAULA mithilfe der folgenden Importer: SpreadsheetImporter. Export ins ANNIS- und PAULA-Format mit dem ANNISExporter und dem PAULAExporter (https://korpling.german.hu-berlin.de/saltnpepper/pepper/download/stable/Pepper_2019.06.11.zip)
Format:	PAULA-XML, ANNIS
Segmentierung der Annotation:	NA basiert auf ‘dipl’

Qualitätsprüfer:	korpling
Datum:	2019
Qualitätsprüfung:	NA
Editor:	Korpling

Typ: Annotationlayer – title

Annotationsart:	Spannenannotation über mehrere norm-Einheiten
Beschreibung:	Es werden alle Werknamen annotiert, auf die ein Autor referiert.

Typ: Annotationvalue – title

Wert:	Wertbeschreibung:
Zeichenkette	Es wird der jeweilige Werkname im Nominativ angegeben. Dies gilt nicht für fremdsprachliches Material. Dieses wird aus dem Original übernommen.

6.3. author_ref

Typ: *Layer* – author_ref

Name:	author_ref
Dokumentation:	Informationen zum gesamten Konvertierungsworkflow gibt es hier: https://github.com/langbank/conversion-pipeline.
Beschreibung:	Inhaltliche Annotationen sind speziell auf das Register des Korpus bezogen. Darunter fallen u.a. Annotationen von im Korpus enthaltenden Begriffen, Definitionen oder Referenzen.

Typ: *Preparationstep* – author_ref

Schritt:	1
Aufbereitung:	Inhaltliche Annotation
Tool	Excel 2013
Format:	Excel 2013

Segmentierung der Annotation:	basiert auf ‘norm’
Qualitätsprüfer:	NA
Datum:	2019
Qualitätsprüfung:	Catharina Fischer
Editor:	Studierende der Humboldt-Universität zu Berlin und der Georg-August-Universität Göttingen

Schritt:	2
Aufbereitung:	Konvertierung von Excel 2013 nach PAULA-XML und ANNIS
Aufbereitungsart:	automatisch
Tool:	Pepper, Konvertierung des Formats Excel nach ANNIS und PAULA mithilfe der folgenden Importer: SpreadsheetImporter. Export ins ANNIS- und PAULA-Format mit dem ANNISExporter und dem PAULAExporter (https://korpling.german.hu-berlin.de/saltnpepper/pepper/download/stable/Pepper_2019.06.11.zip)
Format:	PAULA-XML, ANNIS
Segmentierung der Annotation:	basiert auf ‘norm’
Qualitätsprüfer:	korpling
Datum:	2019
Qualitätsprüfung:	NA
Editor:	Korpling

Typ: *Annotationlayer* – author_ref

Annotationsart:	Spannenannotation über eine oder mehrere norm-Einheiten
Beschreibung:	Markiert Referenzen des Autors auf sich oder andere Autoren bzw. auf beide.

Typ: *Annotationvalue* – author_ref

Wert:	Wertbeschreibung:
pron1pl	Referenz auf den Autor (und auf andere Autoren oder den

	Leser) mit der ersten Person Plural, z.B. „wir“, „unser“.
pron1sg	Referenz auf den Autor mit der ersten Person Singular, z.B. „ich“, „mein“.
pron3sg	Referenz auf einen Autoren mit der dritten Person Singular, z.B. „er“.
pron2pl	Referenz auf mehrere Autoren mit der zweiten Person Plural, z.B. „ihnen“.
author	Explizite Referenz auf einen benannten Autor, z.B. „Plinius“, „Dioscorides“,.“ein Mönch“ (wenn dieser als Autor/Quelle zitiert wird)

6.4. reader_ref

Typ: *Layer* – reader_ref

Name:	reader_ref
Dokumentation:	Informationen zum gesamten Konvertierungsworkflow gibt es hier: https://github.com/langbank/conversion-pipeline.
Beschreibung:	Inhaltliche Annotationen sind speziell auf das Register des Korpus bezogen. Darunter fallen u. a. Annotationen von im Korpus enthaltenden Begriffen, Definitionen oder Referenzen.

Typ: *Preparationstep* – reader_ref

Schritt:	1
Aufbereitung:	Inhaltiche Annotation
Tool	Excel 2013
Format:	Excel 2013
Segmentierung der Annotation:	basiert auf ‘norm’
Qualitätsprüfer:	NA
Datum:	2019
Qualitätsprüfung:	Catharina Fischer
Editor:	Studierende der Humboldt-Universität zu Berlin und der Georg-August-Universität Göttingen

Schritt:	2
Aufbereitung:	Konvertierung von Excel 2013 nach PAULA-XML und ANNIS

Aufbereitungsart:	automatisch
Tool:	Pepper, Konvertierung des Formats Excel nach ANNIS und PAULA mithilfe der folgenden Importer: SpreadsheetImporter. Export ins ANNIS- und PAULA-Format mit dem ANNISExporter und dem PAULAExporter (https://korpling.german.hu-berlin.de/saltnpepper/pepper/download/stable/Pepper_2019.06.11.zip)
Format:	PAULA-XML, ANNIS
Segmentierung der Annotation:	basiert auf 'norm'
Qualitätsprüfer:	korpling
Datum:	2019
Qualitätsprüfung:	NA
Editor:	Korpling

Typ: *Annotationlayer* – reader_ref

Annotationsart:	Spannenannotation über eine oder mehrere norm-Einheiten
Beschreibung:	Markiert Referenzen des Autors auf den Leser.

Typ: *Annotationvalue* – reader_ref

Wert:	Wertbeschreibung:
pron1pl	Referenz mit der ersten Person Plural, also z.B. „wir", „uns", „unser"
pron2sg	Referenz mit der zweiten Person Singular, z. B. „du", „dein"
pron3sg	Referenz mit der dritten Person Singular, z.B. „er", „sie"
pron2pl	Referenz mit der zweiten Person Plural, z.B. „ihr", „euer"
reader	Adressierung an den Leser in nicht-pronominaler Form, z.B. „an den Leser", „lieber Leser"
author	Anrede eines anderen Autoren mit Namen, z.B. *An den Großzachtbahren vnd Hochgelehrten Herrn Johan Jacob Krafften*

Referenzen auf ein generisches „man" oder Imperative werden in dieser Annotationsebene nicht annotiert.

6.5. Metadaten

Dokumentation:	Informationen zum gesamten Konvertierungsworkflow gibt es hier: https://github.com/langbank/conversion-pipeline.
Beschreibung:	Die Metadaten enthalten Informationen in Bezug auf jeden historischen Text im Korpus. Zum Beispiel werden für jeden historischen Text der Titel, der Autor, das Erscheinungsjahr, der Sprachraum und die Textgestaltung angegeben.

Typ: *Preparationstep* - Metadaten

Schritt:	1
Aufbereitung:	Metadatum
Tool	Excel 2013
Format:	Excel 2013
Qualitätsprüfer:	korpling
Datum:	2019
Qualitätsprüfung:	manuell
Editor:	korpling

Schritt:	2
Aufbereitung:	Konvertierung von Excel 2013 nach PAULA-XML und ANNIS
Aufbereitungsart:	automatisch
Tool:	Pepper, Konvertierung des Formats Excel nach ANNIS und PAULA mithilfe der folgenden Importer: SpreadsheetImporter. Export ins ANNIS- und PAULA-Format mit dem ANNISExporter und dem PAULAExporter (https://korpling.german.hu-berlin.de/saltnpepper/pepper/download/stable/Pepper_2019.06.11.zip)
Format:	PAULA-XML, ANNIS
Qualitätsprüfer:	korpling
Datum:	2019
Qualitätsprüfung:	NA
Editor:	Korpling

Metadatum	Annotationswerte	Beschreibung
author	Zeichenkette	Name des Autors des Textes.
bibl	Zeichenkette	Voller bibliographischer Eintrag für die Quelle inklusive der im Korpus annotatierten Seitenzahlen.
annis:doc	Zeichenkette	Annis Dokumentenkürzel (z.B PflantzGart-VR_1639_Rhagor.)
date	JJJJ (Int)	Datum der Veröffentlichung, normalerweise das Jahr.
publisher	Zeichenkette	Verlag/Verleger.
place	Zeichenkette	Verlagsort.
title	Zeichenkette	Titel der Arbeit, aus dem das Dokument extrahiert wurde.
translator	Zeichenkette	Übersetzer des Textes, wenn vorhanden.
trans_from	it lat	Sprache, aus der der Text übersetzt wurde.
editor	Zeichenkette	Herausgeber des Textes (wenn zusätzlich zum Verlag/Verleger angegeben).
version	1.0 2.0 3.0 4.0 5.0 6.0 7.0 8.0 9.0	Version, in der das jeweilige Dokument zum Korpus hinzugefügt wurde.
edition_first	yes no	Wenn bekannt, wird hier angegeben, ob es sich bei dem Faksimile um eine Erstauflage (yes) handelt oder nicht (no).
issue	Int	Nummer des Bandes, wenn vorhanden.
maintopic	science non-science	Angabe darüber, ob der Text eher Alltags- oder Wissenschaftsthemen behandelt.
topic	additiv aus: Al As B G K L M R	Ordnet einem Text einen oder mehrere Themenbereiche zu. Additiver Wert aus den verschiedenen Kürzeln, alphabetisch geordnet. Al: Alchemie, As: Astronomie, B: Botanik, G: Gärtnerei, K: Küche, L: Linguistik, M: Medizin, R: Religion,. Beispielwerte sind "B", "BM" oder "BKM".
register	herbology	Register des Textes.
lingualism	monoling multiling	Gibt an, ob größere Textstellen vorhanden sind, die nicht auf Deutsch verfasst wurden. Dies gilt

		nicht für einzelne Übersetzungen von Pflanzennamen.
orig_date	JJJJ (Int)	Ist ein Text eine Nichterstauflage, dann kann hier das Jahr der Erstauflage stehen, wenn bekannt.
repository	URL	URL zum Repositorium des Faksimiles.
lang_type	enhg nhg	Angabe des Sprachtyps. enhg: frühneuhochdeutsch (early new high german), nhg: neuhochdeutsch (new high german)
lang_area	md obd NA	Das Sprachgebiet wird angegeben. md: mitteldeutsch, obd: oberdeutsch. Bei jüngeren Texten, die schon weitesgtehend standardisiert sind und denen kein Sprachgebiet zugeordnet werden konnte, wird als Wert "NA" angegeben.
text_type	prose lyric mixed	Angaben über die allgemeine Textgestaltung. Prosa: der Text ist komplett in Prosaform verfasst, Poesie: der Text ist komplett in Reimform verfasst; gemischt: der Text ist teilweise prosaisch, teilweise gereimt verfasst.
lyric_type	end_rhyme meter rhyme_meter	Wenn im Metadatum "text_type" die Werte "lyric" oder "mixed" angegeben sind, dann wird hier eingetragen, welche Gestaltungselemente dafür verwendet wurden.
wormwood	yes no	Gibt an, ob ein Textabschnitt zum Thema "Wermut" transkribiert wurde.
herb_sorting	yes no	Gibt an, ob es sich bei dem Text um eine Kräutermonographiesammlung handelt, d.h. dass nacheinander verschiedene Kräuter abgehandelt werden.

7. Referenzen

Bollmann, Marcel; Petran, Florian & Stefanie Dipper (2011): Rule-Based Normalization of Historical Texts. In: Proceedingsof the RANLP Workshop on Language Technologies for Digital Humanities and Cultural Heritage, Hissar, Bulgaria, 34-42.

Dickinson, Markus & Detmar Meurers (2003): Detecting Errors in Part-of-Speech Annotation. In: Proceedings of the 11th Conference of the European Chapter of the Association for Computational Linguistics (EACL-03). Budapest, Hungary.

Druskat, Stephan & Gast, Volker & Krause, Thomas et al. (2016): corpus-tools.org: An Interoperable Generic Software Tool Set for Multi-layer Linguistic Corpora. Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC 2016). http://www.lrec-conf.org/proceedings/lrec2016/summaries/918.html

Krause, Thomas & Amir Zeldes (2016). A new architecture for generic corpus query and visualization. Digital Scholarship in the Humanities, 31(1): 118–139. DOI: https://academic.oup.com/dsh/article/31/1/118/2605493

Krause, Thomas; Lüdeling, Anke; Odebrecht, Carolin & Amir Zeldes (2012): Multiple Tokenizations in a Diachronic Corpus. Exploring Ancient Languages through Corpora Conference (EALC), 14.-16.Juni 2012. Oslo, Norway.

Odebrecht, Carolin; Belz, Malte; Zeldes, Amir; Lüdeling, Anke; Krause, Thomas (2017): RIDGES Herbology: designing a diachronic multi-layer corpus. In: Lang Resources & Evaluation 51 (3), S. 695-725. DOI: https://link.springer.com/article/10.1007/s10579-016-9374-3

Reichmann, Oskar & Klaus-Peter Wegera (Hrsg.) (1993): Frühneuhochdeutsche Grammatik. Tübingen: Niemeyer Verlag.

Schmid, Helmut (1994): Probabilistic Part-of-Speech Tagging Using Decision Trees. Proceedings of International Conference on New Methods in Language Processing, Manchester, UK.

Empfohlene Literatur:

Georges, Karl Ernst (2010) : Ausführliches deutsch-lateinisches Handwörterbuch. Unveränderter Nachdruck der achten verbesserten und vermehrten Auflage von Heinrich Georges. Darmstadt: WBG.

Hartweg Frederic & Klaus-Peter Wegera (2005): Frühneuhochdeutsch. Eine Einführung in die deutsche Sprache des Spätmittelalters und der frühen Neuzeit. 2., neu bearbeitete Auflage, Tübingen: Niemeyer Verlag.

Niermeyer, J.F. & C. van de Kieft (2002): Mediae Latinitatis Lexicon Minus. Überarbeitet von J.W.J. Burgers. Darmstadt: WBG.

ORC: http://www.cis.uni-muenchen.de/ocrworkshop/program.html
Adelung: http://woerterbuchnetz.de/Adelung/
Bibel: http://www.die-bibel.de/online-bibeln
Grimm: http://woerterbuchnetz.de/DWB/
Lexer: http://woerterbuchnetz.de/Lexer/

Anhang

Zuordnung von Annotationen und Dokumenten
Ridges hat eine lange Versionsgeschichte. Deshalb sind nicht alle Annotationen in allen Dokumenten enthalten. Die meisten Annotationsebenen wurden von Studierenden der Humboldt-Universität zu Berlin erstellt. Viele wurden aber auch von einzelnen ForscherInnen angelegt, die selbst entschieden haben, welche Dokumente sie annotieren wollen.

In der folgenden Tabelle sind je Annotationsebene die annotierten Dokumente aufgeführt. Einige Annotationsebenen wurden gruppiert, wenn sie in exakt den gleichen Dokumenten annotiert wurden.

Annotationsebene	annotierte Dokumente
author_ref	AlchymistischePractic_1603_Libavius, AlchymistischePractic-VR_1603_Libavius, ArtzneyBuchleinDerKreutter_1532_Tallat, ArtzneyBuchleinDerKreutter-Abrotanum_1532_Tallat, ArtzneyBuchleinDerKreutter-Centaurea_1532_Tallat, ArtzneyBuchleinDerKreutter-VR_1532_Tallat, BeschreibungDerGraeser_c27-31_1769_Schrebers BeschreibungDerGraeser_c37-48_1769_Schrebers, BlackwellischesKraeuterbuch_1750_Blackwell, BuchDerNatur_1482_vonMegenberg, BuchDerNatur-Canamellis_1482_vonMegenberg ContrafaytKreuterbuch_1532_Brunfels, ContrafaytKreuterbuch-CCLVII-CCLXX_1532_Brunfels, ContrafaytKreuterbuch-CCLXXI-CCXC_1532_Brunfels, ContrafaytKreuterbuch-CCXXXVII-CCXLVIII_1532_Brunfels, ContrafaytKreuterbuch-VR_1532_Brunfels, DeutschePflanzennamen_1870_Grassmann, EinleitungZuDerKraeuterkenntnisz_1764_Oeder, Experimenta_1550_Schellenberg, Experimenta-VR_1550_Schellenbeg, Flora-6_1821_Wilbrand, Flora-7_1821_Sieber, FloraDerPreussischenRheinlande_1870_Wirtgen, FloraSaturnizans_1722_Henckel, GartDerGesundheit_1487_vonCuba, GartDerGesundheit-c15-23_1487_Cuba, GartDerGesundheit-c23-33_1487_Cuba, GartDerGesundheit-VR_1487_vonCuba, GemeinnuetzigesKraeuterbuch_1874_Siegmund, GetreueDarstellungUndBeschreibung_1809_Hayne, GrossesIllustriertesKraeuterbuch_1860_Mueller, GrundrissKraeuterkunde_1792_Willdenow, HortulusSanitatis_1609_Uffenbach, HortulusSanitatis-VR_1609_Uffenbach, Kraeuterbuch_1914_Losch, Kraeutterbuch_1609_Carrichter, Kraeutterbuch-VR_1609_Carrichter, MysterivmSigillorvm_1735_Hiebner, NatürlichePflanzenfamilien_1887_Engler, NeuwKreuterbuch_1588_Theodorus, NeuwKreuterbuch-VR_1588_Theodorus, NewKreueterbuch_1563_Handsch, NewKreuetterBuch_1539_Bock, NewKreuetterbuch-CXXXIIIJ-

	CXXXVIIJ_1539_Bock, NewKreuetterBuch-VR_1539_Bock, NewKreuterbuch_1543_Fuchs, NewKreuterbuch-c19-25_1543_Fuchs, NewKreuterbuch-VR_1543_Fuchs, NewKreuterbuch-VR_1563_Handsch, NochEinigeWorte_1840_Meyen, NochEinigeWorte-VR_1840_Meyen, Paradeiszgaertlein_1588_Rosbach, Paradeiszgaertlein-VR_1588_Rosbach, PflantzGart_1639_Rhagor, PflantzGart-c4_1639_Rhagor, PflantzGart-VR_1639_Rhagor, SchweizerischerBotanicus_1687_vonRoll, SonderbaresKraeuterbuch-21-36_1675_Anonymous, TheatrumBotanicum_1696_Verzascha, ThesaurusSanitatis_1673_Nasser, ThesaurusSanitatis_304-321_1673_Nasser, Unterricht_1774_Eisen, ViridariumReformatum_1719_Valentini, VorlesungenUeberKraeuterkunde_1843_Link, WieSichMeniglich_1557_vonBodenstein, WieSichMeniglich-VR_1557_vonBodenstein, Wund-Artzney_1652_Greiff
clean, dipl, norm, lemma, script, lang, lb, pb, pos, abbr, gram, technicalterm	alle
column	BeschreibungDerGraeser_c27-31_1769_Schrebers BeschreibungDerGraeser_c37-48_1769_Schrebers, BuchDerNatur-Canamellis_1482_vonMegenberg, Experimenta_1550_Schellenberg, Experimenta-VR_1550_Schellenbeg, GartDerGesundheit-c15-23_1487_Cuba, GartDerGesundheit-c23-33_1487_Cuba, HortulusSanitatis-VR_1609_Uffenbach, Kraeutterbuch-VR_1609_Carrichter, NewKreuterbuch-VR_1543_Fuchs, NeuwKreuterbuch_1588_Theodorus, NeuwKreuterbuch-VR_1588_Theodorus, Paradeiszgaertlein-VR_1588_Rosbach
comp, comp_orth, prot	DeutschePflanzennamen_1870_Grassmann, EigenschaftenAllerHeilpflanzen_1828_Anonymous, EinleitungZuDerKraeuterkenntnisz_1764_Oeder, Flora-6_1821_Wilbrand, Flora-7_1821_Sieber, FloraDerPreussischenRheinlande_1870_Wirtgen, FloraSaturnizans_1722_Henckel, GrundrissKraeuterkunde_1792_Willdenow, HortulusSanitatis_1609_Uffenbach, Kraeutterbuch_1609_Carrichter, MysterivmSigillorvm_1735_Hiebner, NochEinigeWorte_1840_Meyen, NochEinigeWorte-VR_1840_Meyen, SchweizerischerBotanicus_1687_vonRoll, SonderbaresKraeuterbuch-1-11_1675_Anonymous, SonderbaresKraeuterbuch-11-21_1675_Anonymous, ThesaurusSanitatis_1673_Nasser, Unterricht_1774_Eisen, VorlesungenUeberKraeuterkunde_1843_Link
comp_n, comp_n_graph, comp_n_mod,	AlchymistischePractic_1603_Libavius, AlchymistischePractic-VR_1603_Libavius, ArtzneyBuchleinDerKreutter_1532_Tallat, ArtzneyBuchleinDerKreutter-Abrotanum_1532_Tallat,

comp_lex, infl_fuge	ArtzneyBuchleinDerKreutter-Cretanus_1532_Tallat, ArtzneyBuchleinDerKreutter-VR_1532_Tallat, BuchDerNatur_1482_vonMegenberg, ContrafaytKreuterbuch_1532_Brunfels, ContrafaytKreuterbuch-CCXXXVII-CCXLVIII_1532_Brunfels, ContrafaytKreuterbuch-CCLXXI-CCXC_1532_Brunfels , ContrafaytKreuterbuch-CCLVII-CCLXX_1532_Brunfels , ContrafaytKreuterbuch-VR_1532_Brunfels, GartDerGesundheit_1487_vonCuba, GartDerGesundheit-c15-23_1487_Cuba, GartDerGesundheit-c23-33_1487_Cuba, GartDerGesundheit-VR_1487_vonCuba, HortulusSanitatis_1609_Uffenbach, Kraeutterbuch_1609_Carrichter, NewKreueterbuch_1563_Handsch, NewKreuetterBuch_1539_Bock, NewKreuetterbuch-CXXXIIIJ-CXXXVIIJ_1539_Bock, NewKreuetterBuch-VR_1539_Bock, NewKreuterbuch_1543_Fuchs, NewKreuterbuch-c19-25_1543_Fuchs, Paradeiszgaertlein_1588_Rosbach, PflantzGart_1639_Rhagor, PflantzGart-c4_1639_Rhagor, PflantzGart-VR_1639_Rhagor, WieSichMeniglich_1557_vonBodenstein, WieSichMeniglich-VR_1557_vonBodenstein, Wund-Artzney_1652_Greiff
comp_amb	AlchymistischePractic_1603_Libavius, ArtzneyBuchleinDerKreutter_1532_Tallat, ArtzneyBuchleinDerKreutter-Abrotanum_1532_Tallat, ArtzneyBuchleinDerKreutter-Cretanus_1532_Tallat, BuchDerNatur_1482_vonMegenberg, ContrafaytKreuterbuch_1532_Brunfels, ContrafaytKreuterbuch-CCLXXI-CCXC_1532_Brunfels, ContrafaytKreuterbuch-CCLVII-CCLXX_1532_Brunfels, GartDerGesundheit_1487_vonCuba, GartDerGesundheit-c15-23_1487_Cuba, GartDerGesundheit-c23-33_1487_Cuba, HortulusSanitatis_1609_Uffenbach, Kraeutterbuch_1609_Carrichter, NewKreueterbuch_1563_Handsch, NewKreuetterBuch_1539_Bock, NewKreuetterBuch-VR_1539_Bock, NewKreuetterbuch-CXXXIIIJ-CXXXVIIJ_1539_Bock, NewKreuterbuch_1543_Fuchs, NewKreuterbuch-c19-25_1543_Fuchs, Paradeiszgaertlein_1588_Rosbach, PflantzGart_1639_Rhagor, PflantzGart-c4_1639_Rhagor, WieSichMeniglich_1557_vonBodenstein, Wund-Artzney_1652_Greiff
comp_a, comp_a_graph	AlchymistischePractic_1603_Libavius, ArtzneyBuchleinDerKreutter_1532_Tallat, ArtzneyBuchleinDerKreutter-VR_1532_Tallat, ArtzneyBuchleinDerKreutter-Abrotanum_1532_Tallat,BuchDerNatur_1482_vonMegenberg, ContrafaytKreuterbuch_1532_Brunfels, ContrafaytKreuterbuch-CCXXXVII-CCXLVIII_1532_Brunfels, ContrafaytKreuterbuch-CCLXXI-CCXC_1532_Brunfels, ContrafaytKreuterbuch-CCLVII-CCLXX_1532_Brunfels, ContrafaytKreuterbuch-VR_1532_Brunfels, GartDerGesundheit_1487_vonCuba, GartDerGesundheit-c15-23_1487_Cuba, GartDerGesundheit-VR_1487_vonCuba,

	HortulusSanitatis_1609_Uffenbach, Kraeutterbuch_1609_Carrichter, NewKreueterbuch_1563_Handsch, NewKreuetterBuch_1539_Bock, NewKreuetterbuch-CXXXIIIJ-CXXXVIIJ_1539_Bock, NewKreuetterBuch-VR_1539_Bock, NewKreuterbuch_1543_Fuchs, NewKreuterbuch-c19-25_1543_Fuchs, Paradeiszgaertlein_1588_Rosbach, PflantzGart_1639_Rhagor, PflantzGart-c4_1639_Rhagor, PflantzGart-VR_1639_Rhagor, WieSichMeniglich_1557_vonBodenstein, WieSichMeniglich-VR_1557_vonBodenstein, Wund-Artzney_1652_Greiff
figure	AlchymistischePractic_1603_Libavius, ArtzneyBuchleinDerKreutter-VR_1532_Tallat, BeschreibungDerGraeser_c27-31_1769_Schrebers BeschreibungDerGraeser_c37-48_1769_Schrebers, BuchDerNatur-Canamellis_1482_vonMegenberg, ContrafaytKreuterbuch_1532_Brunfels, ContrafaytKreuterbuch-CCLVII-CCLXX_1532_Brunfels, ContrafaytKreuterbuch-CCLXXI-CCXC_1532_Brunfels, ContrafaytKreuterbuch-CCXXXVII-CCXLVIII_1532_Brunfels, ContrafaytKreuterbuch-VR_1532_Brunfels, Experimenta_1550_Schellenberg, Experimenta-VR_1550_Schellenbeg, FloraDerPreussischenRheinlande_1870_Wirtgen, GartDerGesundheit_1487_vonCuba, GartDerGesundheit-c15-23_1487_Cuba, GartDerGesundheit-c23-33_1487_Cuba, GemeinnuetzigesKraeuterbuch_1874_Siegmund, GetreueDarstellungUndBeschreibung_1809_Hayne, GrossesIllustriertesKraeuterbuch_1860_Mueller, HortulusSanitatis-VR_1609_Uffenbach, Kraeutterbuch-VR_1609_Carrichter, Kraeuterbuch_1914_Losch, NatürlichePflanzenfamilien_1887_Engler, NewKreuterbuch_1543_Fuchs, NewKreuterbuch-c19-25_1543_Fuchs, NewKreuterbuch-VR_1543_Fuchs, NewKreuterbuch-VR_1563_Handsch, NeuwKreuterbuch_1588_Theodorus, NeuwKreuterbuch-VR_1588_Theodorus, NewKreueterbuch_1563_Handsch, NewKreuetterBuch-VR_1539_Bock, Paradeiszgaertlein_1588_Rosbach, Paradeiszgaertlein-VR_1588_Rosbach, PflantzGart_1639_Rhagor, Phythologia_1662_Becher, SonderbaresKraeuterbuch-1-11_1675_Anonymous, SonderbaresKraeuterbuch-11-21_1675_Anonymous, SonderbaresKraeuterbuch-21-36_1675_Anonymous, TheatrumBotanicum_1696_Verzascha, ViridariumReformatum_1719_Valentini
figure_p	BeschreibungDerGraeser_c27-31_1769_Schrebers BeschreibungDerGraeser_c37-48_1769_Schrebers, BuchDerNatur-Canamellis_1482_vonMegenberg, ContrafaytKreuterbuch-CCLVII-CCLXX_1532_Brunfels, ContrafaytKreuterbuch-CCLXXI-CCXC_1532_Brunfels, Experimenta_1550_Schellenberg, Experimenta-VR_1550_Schellenbeg, HortulusSanitatis-VR_1609_Uffenbach, Kraeutterbuch-VR_1609_Carrichter, NewKreuterbuch-VR_1543_Fuchs, NeuwKreuterbuch_1588_Theodorus, NeuwKreuterbuch-VR_1588_Theodorus, Paradeiszgaertlein-VR_1588_Rosbach,

	Paradeiszgaertlein-VR_1588_Rosbach, SonderbaresKraeuterbuch-21-36_1675_Anonymous
head	AlchymistischePractic_1603_Libavius, AlchymistischePractic-VR_1603_Libavius, Apothekerlexikon_1793_Hahnemann, ArtzneyBuchleinDerKreutter_1532_Tallat, ArtzneyBuchleinDerKreutter-Abrotanum_1532_Tallat, ArtzneyBuchleinDerKreutter-Centaurea_1532_Tallat, ArtzneyBuchleinDerKreutter-VR_1532_Tallat, BeschreibungDerGraeser_c27-31_1769_Schrebers BeschreibungDerGraeser_c37-48_1769_Schrebers, BlackwellischesKraeuterbuch_1750_Blackwell, BuchDerNatur_1482_vonMegenberg, BuchDerNatur-Canamellis_1482_vonMegenberg, ContrafaytKreuterbuch_1532_Brunfels, ContrafaytKreuterbuch-CCLVII-CCLXX_1532_Brunfels, ContrafaytKreuterbuch-CCLXXI-CCXC_1532_Brunfels, ContrafaytKreuterbuch-VR_1532_Brunfels, DeutschePflanzennamen_1870_Grassmann, EigenschaftenAllerHeilpflanzen-149-165_1828_Anonymous, EinleitungZuDerKraeuterkenntnisz_1764_Oeder, Experimenta_1550_Schellenberg, Experimenta-VR_1550_Schellenbeg, Flora-6_1821_Wilbrand, Flora-7_1821_Sieber, FloraDerPreussischenRheinlande_1870_Wirtgen, FloraSaturnizans_1722_Henckel, GartDerGesundheit_1487_vonCuba, GartDerGesundheit-c15-23_1487_Cuba, GartDerGesundheit-c23-33_1487_Cuba, GartDerGesundheit-VR_1487_vonCuba, GetreueDarstellungUndBeschreibung_1809_Hayne, GrossesIllustriertesKraeuterbuch_1860_Mueller, GrundrissKraeuterkunde_1792_Willdenow, HortulusSanitatis_1609_Uffenbach, HortulusSanitatis-VR_1609_Uffenbach, Kraeutterbuch_1609_Carrichter, Kraeutterbuch-VR_1609_Carrichter, MysterivmSigillorvm_1735_Hiebner, NatürlichePflanzenfamilien_1887_Engler, NewKreuterbuch-VR_1543_Fuchs, NeuwKreuterbuch_1588_Theodorus, NeuwKreuterbuch-VR_1588_Theodorus, NewKreueterbuch_1563_Handsch, NewKreuetterBuch_1539_Bock, NewKreuetterbuch-CXXXIIIJ-CXXXVIIJ_1539_Bock, NewKreuetterBuch-VR_1539_Bock, NewKreuterbuch_1543_Fuchs, NewKreuterbuch-c19-25_1543_Fuchs, NewKreuterbuch-VR_1563_Handsch, NochEinigeWorte-VR_1840_Meyen, Paradeiszgaertlein_1588_Rosbach, Paradeiszgaertlein-VR_1588_Rosbach, PflantzGart_1639_Rhagor, PflantzGart-c4_1639_Rhagor, PflantzGart-VR_1639_Rhagor, Phythologia_1662_Becher, SchweizerischerBotanicus_1687_vonRoll, SonderbaresKraeuterbuch-1-11_1675_Anonymous, SonderbaresKraeuterbuch-11-21_1675_Anonymous, SonderbaresKraeuterbuch-21-36_1675_Anonymous, TheatrumBotanicum_1696_Verzascha,

	ThesaurusSanitatis_1673_Nasser, ThesaurusSanitatis_304-321_1673_Nasser, Unterricht_1774_Eisen, ViridariumReformatum_1719_Valentini, VorlesungenUeberKraeuterkunde_1843_Link, WieSichMeniglich_1557_vonBodenstein, WieSichMeniglich-VR_1557_vonBodenstein, Wund-Artzney_1652_Greiff
hi_rend	AlchymistischePractic_1603_Libavius, Apothekerlexikon_1793_Hahnemann, ArtzneyBuchleinDerKreutter_1532_Tallat, ArtzneyBuchleinDerKreutter-Abrotanum_1532_Tallat, ArtzneyBuchleinDerKreutter-Centaurea_1532_Tallat, ArtzneyBuchleinDerKreutter-VR_1532_Tallat, BeschreibungDerGraeser_c27-31_1769_Schrebers BeschreibungDerGraeser_c37-48_1769_Schrebers, BlackwellischesKraeuterbuch_1750_Blackwell, BuchDerNatur_1482_vonMegenberg, BuchDerNatur-Canamellis_1482_vonMegenberg, ContrafaytKreuterbuch_1532_Brunfels, ContrafaytKreuterbuch-CCLVII-CCLXX_1532_Brunfels, ContrafaytKreuterbuch-CCLXXI-CCXC_1532_Brunfels, ContrafaytKreuterbuch-CCXXXVII-CCXLVIII_1532_Brunfels, DeutschePflanzennamen_1870_Grassmann, EigenschaftenAllerHeilpflanzen_1828_Anonymous, EigenschaftenAllerHeilpflanzen-149-165_1828_Anonymous, EinleitungZuDerKraeuterkenntnisz_1764_Oeder, Experimenta_1550_Schellenberg, Experimenta-VR_1550_Schellenbeg, FloraDerPreussischenRheinlande_1870_Wirtgen, FloraSaturnizans_1722_Henckel, GartDerGesundheit_1487_vonCuba, GartDerGesundheit-c15-23_1487_Cuba, GartDerGesundheit-c23-33_1487_Cuba, GemeinnuetzigesKraeuterbuch_1874_Siegmund, GetreueDarstellungUndBeschreibung_1809_Hayne, GrossesIllustriertesKraeuterbuch_1860_Mueller, GrundrissKraeuterkunde_1792_Willdenow, HortulusSanitatis_1609_Uffenbach, HortulusSanitatis-VR_1609_Uffenbach, Kraeutterbuch_1609_Carrichter, Kraeutterbuch-VR_1609_Carrichter, Kraeuterbuch_1914_Losch, MysterivmSigillorvm_1735_Hiebner, NatürlichePflanzenfamilien_1887_Engler, NewKreuterbuch-VR_1543_Fuchs, NeuwKreuterbuch_1588_Theodorus, NeuwKreuterbuch-VR_1588_Theodorus, NewKreueterbuch_1563_Handsch, NewKreuetterBuch_1539_Bock, NewKreuetterbuch-CXXXIIIJ-CXXXVIIJ_1539_Bock, NewKreuetterBuch-VR_1539_Bock, NewKreuterbuch_1543_Fuchs, NewKreuterbuch-c19-25_1543_Fuchs, NewKreuterbuch-VR_1563_Handsch, NochEinigeWorte_1840_Meyen, NochEinigeWorte-VR_1840_Meyen, Paradeiszgaertlein_1588_Rosbach, Paradeiszgaertlein-VR_1588_Rosbach, PflantzGart_1639_Rhagor, PflantzGart-

	c4_1639_Rhagor, PflantzGart-VR_1639_Rhagor, Phythologia_1662_Becher, SchweizerischerBotanicus_1687_vonRoll, SonderbaresKraeuterbuch-1-11_1675_Anonymous, SonderbaresKraeuterbuch-11-21_1675_Anonymous, SonderbaresKraeuterbuch-21-36_1675_Anonymous, TheatrumBotanicum_1696_Verzascha, ThesaurusSanitatis_1673_Nasser, ThesaurusSanitatis_304-321_1673_Nasser, Unterricht_1774_Eisen, ViridariumReformatum_1719_Valentini, VorlesungenUeberKraeuterkunde_1843_Link, WieSichMeniglich-VR_1557_vonBodenstein, Wund-Artzney_1652_Greiff
kaus1, kaus2, kaus3, kaus4, kon1, kon2, lin1, lin2, lin3, lin4, hypo	ArtzneyBuchleinDerKreutter-Abrotanum_1532_Tallat, ArtzneyBuchleinDerKreutter-VR_1532_Tallat, GartDerGesundheit-c23-33_1487_Cuba, GartDerGesundheit-VR_1487_vonCuba, PflantzGart_1639_Rhagor, PflantzGart-VR_1639_Rhagor
morph_ellipsis	AlchymistischePractic_1603_Libavius, BuchDerNatur_1482_vonMegenberg, ContrafaytKreuterbuch_1532_Brunfels, ContrafaytKreuterbuch-CCXXXVII-CCXLVIII_1532_Brunfels, DeutschePflanzennamen_1870_Grassmann, EigenschaftenAllerHeilpflanzen_1828_Anonymous, EinleitungZuDerKraeuterkenntnisz_1764_Oeder, Flora-7_1821_Sieber, FloraDerPreussischenRheinlande_1870_Wirtgen, FloraSaturnizans_1722_Henckel, GrundrissKraeuterkunde_1792_Willdenow, HortulusSanitatis_1609_Uffenbach, Kraeutterbuch_1609_Carrichter, MysterivmSigillorvm_1735_Hiebner, NewKreueterbuch_1563_Handsch, NewKreuetterBuch_1539_Bock, NewKreuetterBuch-VR_1539_Bock, NewKreuterbuch_1543_Fuchs, NochEinigeWorte_1840_Meyen, Paradeiszgaertlein_1588_Rosbach, PflantzGart_1639_Rhagor, PflantzGart-c4_1639_Rhagor, PflantzGart-VR_1639_Rhagor, SchweizerischerBotanicus_1687_vonRoll, SonderbaresKraeuterbuch-1-11_1675_Anonymous, SonderbaresKraeuterbuch-11-21_1675_Anonymous, ThesaurusSanitatis_1673_Nasser, Unterricht_1774_Eisen, VorlesungenUeberKraeuterkunde_1843_Link, WieSichMeniglich_1557_vonBodenstein, Wund-Artzney_1652_Greiff
note	AlchymistischePractic_1603_Libavius, ArtzneyBuchleinDerKreutter_1532_Tallat, BeschreibungDerGraeser_c27-31_1769_Schrebers BeschreibungDerGraeser_c37-48_1769_Schrebers, BuchDerNatur-Canamellis_1482_vonMegenberg, ContrafaytKreuterbuch-CCLVII-CCLXX_1532_Brunfels, ContrafaytKreuterbuch-CCLXXI-CCXC_1532_Brunfels, ContrafaytKreuterbuch-VR_1532_Brunfels, Experimenta_1550_Schellenberg, Experimenta-VR_1550_Schellenbeg, Flora-6_1821_Wilbrand,

	FloraDerPreussischenRheinlande_1870_Wirtgen, FloraSaturnizans_1722_Henckel, HortulusSanitatis_1609_Uffenbach, HortulusSanitatis-VR_1609_Uffenbach, Kraeutterbuch-VR_1609_Carrichter, Kraeuterbuch_1914_Losch, MysterivmSigillorvm_1735_Hiebner, NewKreuterbuch-VR_1543_Fuchs, NewKreuterbuch-VR_1563_Handsch NeuwKreuterbuch_1588_Theodorus, NeuwKreuterbuch-VR_1588_Theodorus, NewKreuetterBuch_1539_Bock, NewKreuetterbuch-CXXXIIIJ-CXXXVIIJ_1539_Bock, NewKreuetterBuch-VR_1539_Bock, NewKreuterbuch-c19-25_1543_Fuchs, NochEinigeWorte_1840_Meyen, NochEinigeWorte-VR_1840_Meyen, Paradeiszgaertlein-VR_1588_Rosbach, PflantzGart_1639_Rhagor, PflantzGart-c4_1639_Rhagor, PflantzGart-VR_1639_Rhagor, TheatrumBotanicum_1696_Verzascha, VorlesungenUeberKraeuterkunde_1843_Link, Wund-Artzney_1652_Greiff
pb_n	AlchymistischePractic_1603_Libavius, AlchymistischePractic-VR_1603_Libavius, Apothekerlexikon_1793_Hahnemann, BeschreibungDerGraeser_c27-31_1769_Schrebers BeschreibungDerGraeser_c37-48_1769_Schrebers, BuchDerNatur-Canamellis_1482_vonMegenberg, ContrafaytKreuterbuch_1532_Brunfels, ContrafaytKreuterbuch-CCLVII-CCLXX_1532_Brunfels, ContrafaytKreuterbuch-CCLXXI-CCXC_1532_Brunfels, ContrafaytKreuterbuch-CCXXXVII-CCXLVIII_1532_Brunfels, ContrafaytKreuterbuch-VR_1532_Brunfels, DeutschePflanzennamen_1870_Grassmann, EigenschaftenAllerHeilpflanzen_1828_Anonymous, EigenschaftenAllerHeilpflanzen-149-165_1828_Anonymous, EinleitungZuDerKraeuterkenntnisz_1764_Oeder, Flora-6_1821_Wilbrand, Flora-7_1821_Sieber, Experimenta_1550_Schellenberg, Experimenta-VR_1550_Schellenbeg, FloraDerPreussischenRheinlande_1870_Wirtgen, FloraSaturnizans_1722_Henckel, GartDerGesundheit-VR_1487_vonCuba, GemeinnuetzigesKraeuterbuch_1874_Siegmund, GrossesIllustriertesKraeuterbuch_1860_Mueller, GrundrissKraeuterkunde_1792_Willdenow, HortulusSanitatis_1609_Uffenbach, HortulusSanitatis-VR_1609_Uffenbach, Kraeuterbuch_1914_Losch, Kraeutterbuch_1609_Carrichter, Kraeutterbuch-VR_1609_Carrichter, MysterivmSigillorvm_1735_Hiebner, NatürlichePflanzenfamilien_1887_Engler, NewKreuterbuch-VR_1543_Fuchs, NewKreuterbuch-VR_1563_Handsch, NeuwKreuterbuch_1588_Theodorus, NeuwKreuterbuch-VR_1588_Theodorus, NewKreueterbuch_1563_Handsch, NewKreuetterBuch_1539_Bock, NewKreuetterbuch-CXXXIIIJ-CXXXVIIJ_1539_Bock, NewKreuetterBuch-VR_1539_Bock, NewKreuterbuch_1543_Fuchs, NochEinigeWorte_1840_Meyen,

	NochEinigeWorte-VR_1840_Meyen, Paradeiszgaertlein_1588_Rosbach, Paradeiszgaertlein-VR_1588_Rosbach, PflantzGart_1639_Rhagor, PflantzGart-c4_1639_Rhagor, PflantzGart-VR_1639_Rhagor, Phythologia_1662_Becher, SchweizerischerBotanicus_1687_vonRoll, SonderbaresKraeuterbuch-1-11_1675_Anonymous, SonderbaresKraeuterbuch-11-21_1675_Anonymous, SonderbaresKraeuterbuch-21-36_1675_Anonymous, TheatrumBotanicum_1696_Verzascha, ThesaurusSanitatis_1673_Nasser, ThesaurusSanitatis_304-321_1673_Nasser, Unterricht_1774_Eisen, ViridariumReformatum_1719_Valentini, VorlesungenUeberKraeuterkunde_1843_Link, WieSichMeniglich_1557_vonBodenstein, Wund-Artzney_1652_Greiff
persname	AlchymistischePractic-VR_1603_Libavius, ArtzneyBuchleinDerKreutter_1532_Tallat, ArtzneyBuchleinDerKreutter-Abrotanum_1532_Tallat, ArtzneyBuchleinDerKreutter-Centaurea_1532_Tallat, ArtzneyBuchleinDerKreutter-VR_1532_Tallat, BeschreibungDerGraeser_c27-31_1769_Schrebers BeschreibungDerGraeser_c37-48_1769_Schrebers, BlackwellischesKraeuterbuch_1750_Blackwell, BuchDerNatur_1482_vonMegenberg, BuchDerNatur-Canamellis_1482_vonMegenberg, ContrafaytKreuterbuch_1532_Brunfels, ContrafaytKreuterbuch-CCLVII-CCLXX_1532_Brunfels, ContrafaytKreuterbuch-CCLXXI-CCXC_1532_Brunfels, ContrafaytKreuterbuch-CCXXXVII-CCXLVIII_1532_Brunfels, ContrafaytKreuterbuch-VR_1532_Brunfels, Experimenta_1550_Schellenberg, Experimenta-VR_1550_Schellenbeg, GartDerGesundheit_1487_vonCuba, GartDerGesundheit-c15-23_1487_Cuba, GartDerGesundheit-c23-33_1487_Cuba, GartDerGesundheit-VR_1487_vonCuba, GemeinnuetzigesKraeuterbuch_1874_Siegmund, GetreueDarstellungUndBeschreibung_1809_Hayne, GrossesIllustriertesKraeuterbuch_1860_Mueller, HortulusSanitatis-VR_1609_Uffenbach, Kraeutterbuch-VR_1609_Carrichter, Kraeuterbuch_1914_Losch, NatürlichePflanzenfamilien_1887_Engler, NewKreuterbuch-VR_1543_Fuchs, NewKreuterbuch-VR_1563_Handsch, NeuwKreuterbuch_1588_Theodorus, NeuwKreuterbuch-VR_1588_Theodorus, NewKreueterbuch_1563_Handsch, NewKreuetterBuch_1539_Bock, NewKreuetterbuch-CXXXIIIJ-CXXXVIIJ_1539_Bock, NewKreuetterBuch-VR_1539_Bock, NewKreuterbuch_1543_Fuchs, NewKreuterbuch-c19-25_1543_Fuchs, Paradeiszgaertlein-VR_1588_Rosbach, SonderbaresKraeuterbuch-21-36_1675_Anonymous, TheatrumBotanicum_1696_Verzascha, ThesaurusSanitatis_304-321_1673_Nasser,

	ViridariumReformatum_1719_Valentini, WieSichMeniglich_1557_vonBodenstein, WieSichMeniglich-VR_1557_vonBodenstein
quote	AlchymistischePractic_1603_Libavius, AlchymistischePractic-VR_1603_Libavius, Apothekerlexikon_1793_Hahnemann, ArtzneyBuchleinDerKreutter-Abrotanum_1532_Tallat, ArtzneyBuchleinDerKreutter-Centaurea_1532_Tallat, ArtzneyBuchleinDerKreutter-VR_1532_Tallat, BeschreibungDerGraeser_c27-31_1769_Schrebers BeschreibungDerGraeser_c37-48_1769_Schrebers, BlackwellischesKraeuterbuch_1750_Blackwell, BuchDerNatur-Canamellis_1482_vonMegenberg, ContrafaytKreuterbuch-CCLVII-CCLXX_1532_Brunfels, ContrafaytKreuterbuch-CCLXXI-CCXC_1532_Brunfels, ContrafaytKreuterbuch-CCXXXVII-CCXLVIII_1532_Brunfels, ContrafaytKreuterbuch-VR_1532_Brunfels, DeutschePflanzennamen_1870_Grassmann, EigenschaftenAllerHeilpflanzen_1828_Anonymous, EigenschaftenAllerHeilpflanzen-149-165_1828_Anonymous, EinleitungZuDerKraeuterkenntnisz_1764_Oeder, Flora-6_1821_Wilbrand, Experimenta_1550_Schellenberg, Experimenta-VR_1550_Schellenbeg, Flora-7_1821_Sieber, FloraDerPreussischenRheinlande_1870_Wirtgen, FloraSaturnizans_1722_Henckel, GartDerGesundheit-c15-23_1487_Cuba, GartDerGesundheit-c23-33_1487_Cuba, GemeinnuetzigesKraeuterbuch_1874_Siegmund, GetreueDarstellungUndBeschreibung_1809_Hayne, GrossesIllustriertesKraeuterbuch_1860_Mueller, GrundrissKraeuterkunde_1792_Willdenow, HortulusSanitatis_1609_Uffenbach, HortulusSanitatis-VR_1609_Uffenbach, Kraeutterbuch-VR_1609_Carrichter, MysterivmSigillorvm_1735_Hiebner, NatürlichePflanzenfamilien_1887_Engler, NewKreuterbuch-VR_1543_Fuchs, NewKreuterbuch-VR_1563_Handsch, NeuwKreuterbuch_1588_Theodorus, NeuwKreuterbuch-VR_1588_Theodorus, NewKreuetterbuch-CXXXIIIJ-CXXXVIIJ_1539_Bock, NewKreuetterBuch-VR_1539_Bock, NewKreuterbuch_1543_Fuchs, NewKreuterbuch-c19-25_1543_Fuchs, NochEinigeWorte_1840_Meyen, NochEinigeWorte-VR_1840_Meyen, Paradeiszgaertlein_1588_Rosbach, Paradeiszgaertlein-VR_1588_Rosbach, PflantzGart-c4_1639_Rhagor, PflantzGart-VR_1639_Rhagor, SonderbaresKraeuterbuch-1-11_1675_Anonymous, SonderbaresKraeuterbuch-11-21_1675_Anonymous, SonderbaresKraeuterbuch-21-36_1675_Anonymous, ThesaurusSanitatis_1673_Nasser, ThesaurusSanitatis_304-321_1673_Nasser, Unterricht_1774_Eisen, VorlesungenUeberKraeuterkunde_1843_Link, WieSichMeniglich_1557_vonBodenstein, WieSichMeniglich-

	VR_1557_vonBodenstein, Wund-Artzney_1652_Greiff
quotemark	HortulusSanitatis-VR_1609_Uffenbach, BeschreibungDerGraeser-c27-31_1769_Schrebers
reader_ref	AlchymistischePractic_1603_Libavius, AlchymistischePractic-VR_1603_Libavius, ArtzneyBuchleinDerKreutter_1532_Tallat, ArtzneyBuchleinDerKreutter-Abrotanum_1532_Tallat, ArtzneyBuchleinDerKreutter-VR_1532_Tallat, BeschreibungDerGraeser_c27-31_1769_Schrebers BeschreibungDerGraeser_c37-48_1769_Schrebers, BuchDerNatur_1482_vonMegenberg, BuchDerNatur-Canamellis_1482_vonMegenberg, ContrafaytKreuterbuch_1532_Brunfels, ContrafaytKreuterbuch-CCLVII-CCLXX_1532_Brunfels, ContrafaytKreuterbuch-CCLXXI-CCXC_1532_Brunfels, ContrafaytKreuterbuch-CCXXXVII-CCXLVIII_1532_Brunfels, EinleitungZuDerKraeuterkenntnisz_1764_Oeder, Experimenta_1550_Schellenberg, Experimenta-VR_1550_Schellenbeg, Flora-6_1821_Wilbrand, FloraSaturnizans_1722_Henckel, GartDerGesundheit_1487_vonCuba, GartDerGesundheit-c15-23_1487_Cuba, GartDerGesundheit-c23-33_1487_Cuba, GartDerGesundheit-VR_1487_vonCuba, GemeinnuetzigesKraeuterbuch_1874_Siegmund, GrossesIllustriertesKraeuterbuch_1860_Mueller, GrundrissKraeuterkunde_1792_Willdenow, HortulusSanitatis-VR_1609_Uffenbach, Kraeutterbuch_1609_Carrichter, Kraeutterbuch-VR_1609_Carrichter, MysterivmSigillorvm_1735_Hiebner, NewKreueterbuch_1563_Handsch, NewKreuetterBuch_1539_Bock, NewKreuetterbuch-CXXXIIIJ-CXXXVIIJ_1539_Bock, NewKreuetterBuch-VR_1539_Bock, NewKreuterbuch_1543_Fuchs, NewKreuterbuch-c19-25_1543_Fuchs, NochEinigeWorte_1840_Meyen, Paradeiszgaertlein_1588_Rosbach, Paradeiszgaertlein-VR_1588_Rosbach, PflantzGart-c4_1639_Rhagor, PflantzGart-VR_1639_Rhagor, SchweizerischerBotanicus_1687_vonRoll, TheatrumBotanicum_1696_Verzascha, ThesaurusSanitatis_1673_Nasser, ThesaurusSanitatis_304-321_1673_Nasser, Unterricht_1774_Eisen, ViridariumReformatum_1719_Valentini, VorlesungenUeberKraeuterkunde_1843_Link, WieSichMeniglich_1557_vonBodenstein, WieSichMeniglich-VR_1557_vonBodenstein, Wund-Artzney_1652_Greiff
title	AlchymistischePractic-VR_1603_Libavius, ArtzneyBuchleinDerKreutter_1532_Tallat, ArtzneyBuchleinDerKreutter-Abrotanum_1532_Tallat, BeschreibungDerGraeser_c27-31_1769_Schrebers BeschreibungDerGraeser_c37-48_1769_Schrebers, BlackwellischesKraeuterbuch_1750_Blackwell, BuchDerNatur-Canamellis_1482_vonMegenberg, ContrafaytKreuterbuch-CCLVII-

	CCLXX_1532_Brunfels, ContrafaytKreuterbuch-CCLXXI-CCXC_1532_Brunfels, ContrafaytKreuterbuch-VR_1532_Brunfels, Experimenta_1550_Schellenberg, Experimenta-VR_1550_Schellenbeg, GartDerGesundheit_1487_vonCuba, GartDerGesundheit-c15-23_1487_Cuba, GartDerGesundheit-c23-33_1487_Cuba, GartDerGesundheit-VR_1487_vonCuba, GetreueDarstellungUndBeschreibung_1809_Hayne, HortulusSanitatis-VR_1609_Uffenbach, Kraeutterbuch-VR_1609_Carrichter, Kraeuterbuch_1914_Losch, NewKreuterbuch-VR_1543_Fuchs, NewKreuterbuch-VR_1563_Handsch, NeuwKreuterbuch_1588_Theodorus, NeuwKreuterbuch-VR_1588_Theodorus, NewKreuetterBuch_1539_Bock, NewKreuetterbuch-CXXXIIIJ-CXXXVIIJ_1539_Bock, NewKreuetterBuch-VR_1539_Bock, NewKreuterbuch-c19-25_1543_Fuchs, Paradeiszgaertlein-VR_1588_Rosbach, Phythologia_1662_Becher, TheatrumBotanicum_1696_Verzascha, ThesaurusSanitatis_304-321_1673_Nasser, ViridariumReformatum_1719_Valentini
unclear	AlchymistischePractic_1603_Libavius, ArtzneyBuchleinDerKreutter_1532_Tallat, ArtzneyBuchleinDerKreutter-Abrotanum_1532_Tallat, BeschreibungDerGraeser_c27-31_1769_Schrebers BeschreibungDerGraeser_c37-48_1769_Schrebers, BlackwellischesKraeuterbuch_1750_Blackwell, BuchDerNatur-Canamellis_1482_vonMegenberg, ContrafaytKreuterbuch_1532_Brunfels, ContrafaytKreuterbuch-CCLVII-CCLXX_1532_Brunfels, DeutschePflanzennamen_1870_Grassmann, EigenschaftenAllerHeilpflanzen_1828_Anonymous, EinleitungZuDerKraeuterkenntnisz_1764_Oeder, Experimenta_1550_Schellenberg, Experimenta-VR_1550_Schellenbeg Flora-6_1821_Wilbrand, Flora-7_1821_Sieber, FloraSaturnizans_1722_Henckel, GartDerGesundheit_1487_vonCuba, GartDerGesundheit-c15-23_1487_Cuba, GartDerGesundheit-VR_1487_vonCuba, GetreueDarstellungUndBeschreibung_1809_Hayne, GrundrissKraeuterkunde_1792_Willdenow, HortulusSanitatis_1609_Uffenbach, HortulusSanitatis-VR_1609_Uffenbach, Kraeutterbuch_1609_Carrichter, Kraeutterbuch-VR_1609_Carrichter, MysterivmSigillorvm_1735_Hiebner, NewKreuterbuch-VR_1543_Fuchs, NewKreuterbuch-VR_1563_Handsch, NeuwKreuterbuch_1588_Theodorus, NeuwKreuterbuch-VR_1588_Theodorus, NewKreuetterBuch_1539_Bock, NewKreuetterBuch-VR_1539_Bock, Paradeiszgaertlein_1588_Rosbach, Paradeiszgaertlein-VR_1588_Rosbach, PflantzGart_1639_Rhagor, PflantzGart-c4_1639_Rhagor, SchweizerischerBotanicus_1687_vonRoll,

	SonderbaresKraeuterbuch-1-11_1675_Anonymous, SonderbaresKraeuterbuch-11-21_1675_Anonymous, SonderbaresKraeuterbuch-21-36_1675_Anonymous, ThesaurusSanitatis_1673_Nasser, Unterricht_1774_Eisen, WieSichMeniglich_1557_vonBodenstein, Wund-Artzney_1652_Greiff

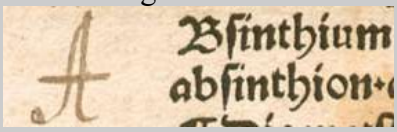
Transkription und Normalisierung

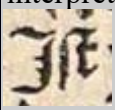
Die Transkription von Faksimiles stellt für die korpuslinguistische Aufbereitung zumeist die grundlegende, diplomatische Ebene (**dipl**). Diese Ebene soll sich grafisch hinsichtlich Orthografie, Getrennt- und Zusammenschreibung und Sonderzeichen möglichst nah am zugrunde liegenden Faksimile orientieren.
Die **clean**-Ebene enthält erste Normalisierungen hinsichtlich Sonderzeichen und grafischer Strukturierungen. So werden u. a. Ligaturen, die nach moderner Rechtschreibung nicht mehr verwendet werden, normalisiert. Graphische Markierungen der einzelnen Texte wie Zeilenumbrüche werden aufgelöst und Sonderzeichen einiger Fonts wie Fraktur auf die heutigen Schriftsätze übertragen.
Die **norm**-Ebene stellt einen weiteren Normalisierungsschritt dar, indem hier die Tokenisierung, die Orthografie, Flexionsformen und Wortbildungen einheitlich dem nhd. Sprachstand angepasst werden.
Durch die einzelnen Normalisierungsschritte und der daraus resultierenden multiplen Segmentierung ist es möglich, die uneinheitliche Orthografie in einzelnen Texten sowie im diachronen Verlauf zu vereinheitlichen und dennoch die spezifischen Besonderheiten des jeweiligen Textes/der jeweiligen Sprachstufe im Korpus abzubilden. Weiterhin können automatische Tagging- und ParsingTools, die in vielen Fällen auf modernem Deutsch trainiert sind, auf der norm-Ebene angewandt werden. Durch die Alignierung der einzelnen Ebenen können Annotationen und Suchanfragen, die auf Grundlage der norm-Ebene getätigt wurden, auf die dipl-Ebene zurückgeführt werden.
Nachfolgend werden die Richtlinien zur Transkription und Normalisierung im Einzelnen gegenübergestellt. Allgemein für jede Segmentierungsebene sowie Annotationsebene gilt, dass keine Leerzeilen oder Leerzeichen enthalten sein dürfen. Dazu wird für jeden Fall ein Beispiel aus dem Ridges Herbology Version 4.0 Korpus gegeben, die konkrete Regel beschrieben und die weiteren Normalisierungsschritte gegenübergestellt (dipl – clean – norm). Die Richtlinien sind nach Schwerpunkten gruppiert: (1) Zeichensetzung/Sonderzeichen (2) Segmentierung/Tokenisierung (3) Interpunktion.

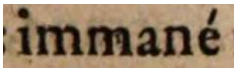
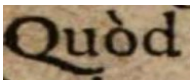
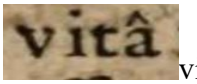
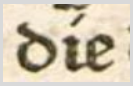
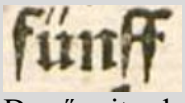
1 Transkribieren
  1.1 Lautzeichen
  1.2 Diakritika
  1.3 Satzzeichen
  1.4 Sonstige Zeichen
  1.5 Ligaturen
  1.6 Griechische Schrift
  1.7 Typographische Besonderheiten
2 Tokenisieren
  2.1 Wort
  2.2 Zeilenumbrüche
  2.3 Abkürzungen
  2.4 Zahlen
  2.5 Satzzeichen
  2.6 Fußnoten
3 Normalisieren
  3.1 Phonologie

3.2 Morphologie
3.3 Morphosyntax
3.4 Wortbildung
3.5 Lexik
3.6 Graphematik

Sonderzeichen wie zum Beispiel das „ſ“ werden mit Hilfe von Unicodes in der dipl-Ebene realisiert. Eine Liste der zu verwendenden Unicodes ist angefügt.

Nr.	**dipl-Ebene**	**clean-Ebene**	**norm-Ebene**
	Transkript des Faksimiles.	**Wird automatisch durch ein Python-Skript (clean-skript.py) erstellt. Dieses ersetzt alle heute unüblichen Sonderzeichen durch heute verwendete Entsprechungen.**	**Erfolgt manuell im .xlsx Format nach der modernen neuen Rechtschreibung. Tipp: Kopieren Sie sich die clean-Ebene und verändern Sie dann die entsprechenden Stellen!**
1.0	Nicht mehr lesbare Zeichen oder Zeichenketten werden mit einem Unterstrich markiert, unabhängig davon, wieviele Zeichen (in etwa) nicht mehr interpretiert werden können.	Durch die fehlende Transkription ist eine Normalisierung nicht mehr möglich, Dies wird mit dem Platzhalter-Tag „unknown“ angezeigt.	Durch die fehlende Transkription ist eine Normalisierung nicht mehr möglich, Dies wird mit dem Platzhalter-Tag „unknown“ angezeigt.
	dipl	**clean**	**norm**
	_	unknown	unknown
	Handschriftliche Versalien die offensichtlich zum Text gehören (KEINE Anmerkungen oder Kommentare von Lesern), werden mit annotiert. 		
	dipl	**clean**	**norm**
	ABſinthium	ABsinthium	Absinthium
1.1	Die Zeichen für 'I' und 'J' sind in Fraktur-Texten graphisch in der Regel nicht voneinander unterscheidbar. Konvention in Ridges	Wie in der dipl-Ebene	Anpassung an moderne Orthographie.

	ist daher, das Zeichen in der dipl als 'J' darzustellen und erst in der norm-Ebene das Zeichen individuell nach moderner Orthografie zu interpretieren. 		
	dipl	**clean**	**norm**
	Jſt	Jst	Ist
1.2	Das „ſ“ wird übernommen.	Das lange „ſ“ wird durch ein reguläres „s“ ersetzt.	
	Dipl	**clean**	**norm**
	Leſer	Leser	Leser
	 Dies ist ein <x>		
	dipl	**clean**	**norm**
	x	x	x
	Das Zeichen 'Ȝ' wird in der dipl-Ebene nicht von 'Z' unterschieden.	Wie in der dipl-Ebene	
	dipl	**clean**	**norm**
	 Dies ist ein <z>	z	z
	In manchen Drucken sieht das Zeichen „Y“ etwas anders aus; es wird aber als „normales Y“ transkribiert.		
	dipl	**clean**	**norm**
	 Dis ist ein <y>	y	y
1.3	Alle Akzente und Diakritika werden	Alle Akzente werden übernommen.	Alle Akzente werden übernommen.

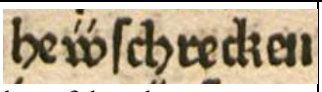
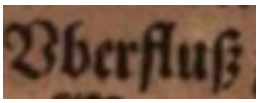
	übernommen.		
	Dipl	**clean**	**norm**
	 immané	immané	immané
	 Quòd	Quòd	Quòd
	 vitâ	vitâ	vitâ
	Dieses Zeichen wird übernommen.	Dieses diakritische Zeichen wird in der clean-Ebene nicht mehr realisiert und durch den zugrundeliegenden Vokal ersetzt.	
	Dipl	**clean**	**norm**
	zů	zu	zu
	Umlaute mit superskribiertem „e“ werden übernommen.	Umlaute mit superskribiertem „e“ werden wie moderne Umlaute des Deutschen realisiert.	
	dipl	**clean**	**norm**
	Glåſer	Gläser	Gläser
	 'í' mit schrägem Strich wird nicht vom ‘i’ mit Punkt unterschieden.		
	dipl	**clean**	**norm**
	die	die	die
	 Das ű mit schrägen Strichen wird nicht nicht vom ü mit Punkten unterschieden.	ü mit schrägen Strichen wird in der clean-Ebene durch 'ü' ersetzt.	ü mit schrägen Strichen wird in der norm-Ebene durch 'ü' ersetzt.

	dipl	clean	norm
	fünff	fünff	fünf
	Alle horizonalen Striche über einem Zeichen werden als Tilde interpretiert Eine Tilde kann stehen für: ausgelassene Nasale, sog. Nasalstrich	Tilden werden entsprechend ihrer Funktion umgesetzt. Die kann u. a. aus Reichmann & Wegera (1993) entnommen werden.[16] Achtung: In der clean-Ebene wird der Strich automatisch als Nasalstrich interpretiert. Gegebenenfalls in der norm ausgleichen.	Gegebenenfalls in der clean-Ebene eingefügte Nasale anders umsetzen.
	dipl	**clean**	**norm**
	om̃ia	omnia	omnia
	sciẽtiã meã	scientiam meam	scientiam meam
	eĩ	ein	ein
	samẽ	samen	Samen
	einẽ	eynem	einem
		nemlich	nämlich
	Verdopplung eines Buchstabens		
	dipl	**clean**	**norm**
	him̃el	himmel	Himmel
		inn	in
	Das Dicit-Zeichen (Abkürzung für “der”) wird in der dipl-Ebene mit ð übernommen	Es wird in der clean automatisch mit "der" ersetzt	In der norm auch.

16 Nachweis für das Beispiel: Reichmann & Wegera 1993: 31.

	dipl	**clean**	**norm**
	oð	oder	oder
	Dieses Zeichen wird mit □ repräsentiert.	Ersetzung durch <us>	Ersetzung durch <us>
	dipl	**clean**	**norm**
	de⁹	deus	deus
		Ersetzung durch <er>	Ersetzung durch <er>
	dipl	**clean**	**norm**
	v̊tzerẽ	vertzeren	verzehren
1.4	Das Zeichen für morphologische Worttrennung (oder Zeilenumbruch) wird mit ⸗ übernommen.	Die morphologische Trennung durch „⸗“ bleibt erhalten, wird jedoch mit einem regulären „-“ ersetzt. Komposita, die in der dipl-Ebene mit einem „-“ realisiert werden, behalten dieses Zeichen bei, um die allgemeine Suche nach Komposita in dieser Ebene zu ermöglichen. Für Erläuterungen zur rein graphischen Trennung, wie Zeilenumbrüchen, siehe Nummer 2.1.	Gleichheitszeichen, die Kompositabildung anzeigen, werden durch Minuszeichen ersetzt oder das Kompositum wird analog zur modernen Orthographie zusammengeschrieben. (Grundlage hierfür ist der aktuelle Stand des Dudens)
	Dipl	**clean**	**norm**
	Dañ⸗reiſz	Dann-reisz	Tannenreis
	Das "modifizierende Pluszeichen" wird übernommen.	In der clean wird das Zeichen zu ":" normalisiert.	In der norm wird das Zeichen zu ":" normalisiert.
	dipl	**clean**	**norm**
	⁘	:	:
	Das Absatzzeichen wird in dipl-, clean- und norm-Ebene übernommen.		

	dipl	**clean**	**norm**
	¶	¶	¶
1.5	Bruchangaben werden mit „/“ übernommen. Dabei wird getrennt segmentiert, um gemischte Brüche transparent zu halten.	Bruchangaben werden mit „/“ übernommen. Dabei wird getrennt segmentiert, um gemischte Brüche transparent zu halten.	Bruchangaben werden mit „/“ übernommen. Dabei wird getrennt segmentiert, um gemischte Brüche transparent zu halten.
	Dipl	**clean**	**norm**
	1	1	1
	1/2	1/2	1/2
	Dieses Symbol wird übernommen.	Dieses Symbol wird übernommen.	Dieses Symbol wird übernommen.
	Dipl	**clean**	**norm**
	⊙	⊙	⊙
	Senkrechtstriche werden mit „|“ übernommen.	Senkrechtstriche werden übernommen.	Senkrechtstriche werden übernommen.
	dipl	**clean**	**norm**
	dieſe	diese	diese
	Dieses Symbol wird als “$” transkribiert. Es bedeutet wahrscheinlich „Zwitter“ bzw. „zwitterig“.	Die als Platzhalter dienenden Dollarzeichen werden übernommen.	Dollarzeichen werden übernommen. In der Ebene „comment“ kann ergänzt werden, dass das Zeichen vermutlich „Zwitter“ oder „zwitterig“ bedeutet.
	dipl	**clean**	**norm**
	$	$	$
	Dieses Zeichen wird als		

	7 transkribiert. Es handelt sich hierbei um das Tironische „et“.		
	7	et	et
1.6	Die Ligaturen „æ“ und „Æ“ werdem übernommen. Das gleiche gilt ebenfalls für „Œ“ „œ“	Die Ligatur wird aufgelöst.	
	dipl	**clean**	**norm**
	hæc	haec	haec
	 Die ct-Ligatur wird nicht übernommen.	Die ct-Ligatur wird nicht übernommen.	Die ct-Ligatur wird nicht übernommen.
	dipl	**clean**	**norm**
	Lactucis	Lactucis	Lactucis
	 Dies ist eine Ligatur aus v̈ und v. Die Ligatur wird aufgelöst v̈v	v̈v	
	dipl	**clean**	**norm**
	 hev̈vſchrecken	hev̈vschrecken	Heuschrecken
	Die ſz - Ligatur wird mit dem langem „ſ“ und „z“ übernommen.	Die Ligatur wird analog zu „sz“ aufgelöst.	Nach modernen Orthografieregeln wird für „sz“ das „ß“ verwendet.
	dipl	**clean**	**norm**
	 Uberflu**ſz**	Uberflusz	Überfluss
	Griechische Ligaturen werden nicht abgebildet, sondern bereits in der dipl		

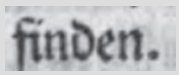
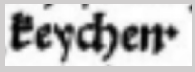
	aufgelöst: Ligatur aus o und ς Als Hilfe: https://de.wikipedia.org/wiki/Griechisches_Alphabet#/media/File:Greek_alphabet_ligatures.jpg		
	dipl	**clean**	**norm**
	ἄκανθος	ἄκανθος	ἄκανθος
	 Dies beteutet “etc.”		
	dipl	**clean**	**norm**
	& c.	& c.	etc.
	Das kaufmännische Und „&“ übernommen.	Das kaufmännische Und wird übernommen.	Das kaufmännische Und wird übernommen.
	dipl	**clean**	**norm**
	&	&	&
1.7	Sprachliches Material mit anderen Schriftsystemen (z.B. Griechisch) wird in clean und norm nicht geändert http://www.unicode.org/charts/PDF/U0370.pdf sowie http://www.unicode.org/charts/PDF/U1F00.pdf		
	dipl	**clean**	**norm**
	ἄκανθος	ἄκανθος	ἄκανθος

1.8	Oft werden die ersten Zeichen bzw. ersten Wörter (im Bsp. Von Weg⸗) einer neuen Seite bei einem Seitenumbruch in der vorhergehenden Seite in einem eigenen Absatz/in einer eigenen Zeile doppelt realisert (sog. Kustoden). Diese Zeichen bzw. das Wort wird nicht mit in das Transkript aufgenommen.		
2.0	Auch entgegen moderner Orthographieregeln wird analog zur Textgrundlage tokenisiert.		Wortbildung und Großschreibung, die nicht der modernen Orthographieregeln entsprechen, werden angeglichen.
	dipl	**clean**	**norm**
	Spannen⸗lang	Spannen-lang	spannenlang
2.1	Klitisierte Elemente können zusammengeschrieben (z.B. *ichs, bedarfs, aufs*) oder apostrophiert dargesetllt sein (z.B. *ich's, bedarf's, auf's*). Klitisiert werden schwach betonte Morpheme. Z.B.: Formen des Personalpronomens, unbestimmter oder bestimmter Artikel	Klitika werden als ein Token realisiert.	Klitika werden in dieser Ebene aufgelöst. Dazu wird die Tokenisierung verändert, in dem aus einem Token in dipl (oder clean) zwei Token gemacht werden!
	dipl	**clean**	**norm**
	ichs	ichs	ich
			es
	Auch entgegen moderner Orthographieregeln wird analog zur Textgrundlage tokenisiert.		Getrennt- und Zusammenschreibung wird nach modernen Orthografieregeln angepasst. Dazu wird die Tokenisierung verändert, indem aus einem Token in der dipl-Ebene (oder clean) zwei Token in der

			norm-Ebene gemacht werden!
	dipl	**clean**	**norm**
	zuſetzen	zusetzen	zu
			setzen
	Im Falle einer Worttrennung aufgrund von Zeilenumbrüchen, die keine overte, graphische Markierung wie „=/-“ beinhalten, werden formal die getrennten Elemente als jeweils ein Token in der dipl-Ebene betrachtet.	Die Worttrennung ohne graphische Markierung ist nicht in jedem Fall transparent. Graphische Normalisierungen werden nicht vorgenommen.	Diese Art der Worttrennung, wenn sie transparent ist, wird nach modernen Orthographieregeln normalisiert und die graphische Trennung aufgehoben.
	dipl	**clean**	**norm**
	ge	ge	genannt
	nent	nent	
	Komposita, egal welcher Komplexität oder Bildungsweise, mit Gleichheitszeichen werden als ein Token realisiert.	Die morphologische Trennung durch „⸗“ bleibt erhalten, wird jedoch mit einem regulären „-“ ersetzt.	Gleichheitszeichen, die Kompositabildung anzeigen, werden durch Minuszeichen ersetzt oder das Kompositum wird analog zur modernen Orthografie zusammengeschrieben. (Grundlage hierfür bildet die aktuelle Ausgabe des Dudens.)
	dipl	**clean**	**norm**
	Artzney⸗Kunſt	Artzney-Kunst	Arzneikunst
2.2	Von Zeilenumbrüchen betroffene Elemente werden analog zur Primärquelle getrennt tokenisiert. Das Trennungszeichen „⸗“	Wörter, die von Zeilenumbrüchen betroffen sind, werden in der clean-Ebene ohne den (Doppel-) Bindestrich, der diesen anzeigt, sowie in einer	

	wird übernommen.	Spanne zusammengefasst realisiert. Die grafische Worttrennung wird somit aufgehoben.	
	dipl	**clean**	**norm**
	allge⸗ meinen	allgemeinen	allgemeinen
	Von Zeilenumbrüchen betroffene Komposita werden analog zur Primärquelle getrennt tokenisiert, dabei bleibt die Kennzeichnung der morphologischen Worttrennung unberührt.	Wenn ein Kompositum, das durch Gleichheitszeichen grundsätzlich getrennt wird, von einem Zeilenumbruch betroffen ist, so wird dieses Gleichheitszeichen in der clean-Ebene entfernt, das andere Gleichheitszeichen für die morphologische Trennung wird analog zu Nummer 1.1 mit einem Minus ersetzt. Das Kompositum wird in einer Spanne zusammengefasst wiedergegeben.	Die Gleichheitszeichen, die Kompositabildung anzeigen, werden durch Minuszeichen ersetzt oder das Kompositum wird analog zur modernen Orthografie zusammengeschrieben. (Grundlage hierfür bildet die aktuelle Ausgabe des Dudens.)
	dipl	**clean**	**norm**
	Stab⸗ wurtz⸗Oel	Stabwurtz-Oel	Stabwurzöl
2.3	Abkürzungen werden analog zur Textgrundlage tokenisiert. Das heißt, auch Setzfehler oder Spatien, die nicht modernen Orthografieregeln entsprechen, werden realisiert.		Abkürzungen werden nach Dudenrichtlinien umgesetzt. Dies hat zur Folge, dass Abkürzungen unterschiedlich tokenisiert werden können (bspw. Abkürzungen als zwei oder mehr Tokens wie u. a. m. vs. solche, die als ein Token realisiert werden, wie usw.). vgl. dazu Nummer 2/3.1 und 2/3.3. In Fällen, in denen keine Dudenrichtlinie zugeordnet werden kann, wird immer nach einem Punkt segmentiert.

	dipl	**clean**	**norm**
	u.	u.	usw.
	ſ.	s.	
	w.	w.	
	Abkürzungen werden analog zur Textgrundlage tokenisiert. Das heißt, auch Setzfehler oder Spatien, die nicht modernen Orthografieregeln entsprechen, werden realisiert.		Abkürzungen werden nach Dudenrichtlinien umgesetzt. Dies hat zur Folge, dass Abkürzungen unterschiedlich tokenisiert werden können (bspw. Abkürzungen als zwei oder mehr Tokens wie u. a. m. vs. solche, die als ein Token realisiert werden, wie usw.). vgl. dazu Nummer 2/3.1 und 2/3.2. In Fällen, in denen keine Dudenrichtlinie zugeordnet werden kann, wird immer nach einem Punkt segmentiert.
	dipl	**clean**	**norm**
	u	u	u.
			d.
	d.g.	d.g.	g.
2.4	Punktsetzungen bei Ordinalzahlen werden mit der Ziffer als ein Token realisiert.		Nach modernen Orthographieregeln werden Ordinalzahlen mit Punkt realisiert.
	dipl	**clean**	**norm**
	I.	I.	I.
	Die Faksimilia weisen oft Kardinalzahlen mit Interpunktion auf. Diese Punktsetzungen werden mit der jeweiligen Ziffer als Token realisiert. Ausnahmen bilden		Um die Kardinalzahlen von Ordinalzahlen in der norm-Ebene getrennt betrachten zu können, wird die Interpunktion bei Kardinalzahlen weggelassen.

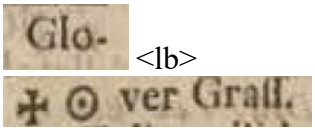
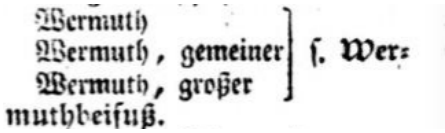
	Zahlen, die mit einem satzbeenden Punkt auftreten. Hier wird wie gewohnt die Satzinterpunktion getrennt von der Kardinalzahl tokenisiert. 		
	dipl	**clean**	**norm**
	I.	I.	I
	Loth	Loth	Lot
	.IX.	.IX.	IX
	Zeichenketten wie "5 %", "5-12", "800'" werden auf der dipl und clean als ein Token und in der norm als mehrere Token betrachtet, wenn sie im Scan visuell zusammen stehen. Stehen sie visuell auseinander, werden sie auch als verschiedene Tokens realisiert.		Norm: immer als mehrere Token
2.5	 Satzinterpunktion wie Kommata, Punkte, Virgeln oder Semikola werden als jeweils eigenständige Token realisiert.		
	dipl	**clean**	**norm**
	finden	finden	finden
	.	.	.
			

	Die Höhe von Punkten wird nicht beachtet. Da wir nicht immer zweifelsfrei feststellen können, wann ein Punkt "hoch" ist und wann "tief" transkribieren wir das alles als normale Punkte.		
	keychen	keychen	keuchen
	.	.	.
	Dieses Zeichen wird als Punkt transkribiert.		
	werden	werden	werden
	.	.	.
2.6	Fußnotenreferenzierungen werden in dipl- und clean-Ebene als mehrere Token und in der norm-Ebene als ein Token realisiert		Norm: ein Token
	dipl	**clean**	**norm**
	(*) * *)	(*) * *)	(*) **)
3.1	Formen die in phonologischer Hinsicht fnhd. oder dialektalen Lautstand aufweisen werden in der norm den (standard)nhd Entsprechungen angepasst		
	dipl	**clean**	**norm**
	Hellenpein	Hellenpein	Höllenpein
	wänlin	wänlin	Wännlein

	Dialektale Formen		
	dipl	**clean**	**norm**
	beede	beede	beide
	Apokope / Synkope (hier: Synkope)		
	dipl	**clean**	**norm**
	hänget	hänget	hängt
3.2	Fnhd. Flexionsformen werden durch die nhd. Entsprechungen ersetzt.		
	dipl	**clean**	**norm**
	in die Nasen das kal Haupt	in die Nasen das kal Haupt	in die Nase das kahle Haupt
	Es erfolgt keine Anpassung des Genus		
	dipl	**clean**	**norm**
	das Milz	das Milz	das Milz
	Vom Nhd. abweichende starke oder schwache Verbalflexion bzw. Rückumlaut wird dem nhd. Stand angepasst		
	gennent		*genannt*
	gebauen		*gebaut*
	Auseinanderschreibung		
	obgenannt		*oben genannt*
	hiebevor		*hier bevor*
3.3	Die aus der niederfränkischen Schreibtradition stammende <ij>-Schreibung für lateinisch <ii> wird in der norm zu <ii> normalisiert.		
	dipl	**clean**	**norm**
	Lapatijs	Lapatijs	Lapatiis
3.4	Ausgestorbenes lexikalisches Material wird mit modernen Orthographieregeln übernommen, aber nicht lexikografisch übersetzt.		
	dipl	**clean**	**norm**
	Vergeſz	Vergeſz	Vergess

<table>
<tr><td rowspan="3">3.5</td><td colspan="3">Es erfolgt keine morphosyntaktische Anpassung (z.B. innerhalb der Nominalphrase)</td></tr>
<tr><td>dipl</td><td>clean</td><td>norm</td></tr>
<tr><td>(es heilt)
die verſehrte
Därmelein</td><td>die verſehrte Därmelein</td><td>die verſehrte Därmelein</td></tr>
<tr><td rowspan="3">3.6</td><td colspan="3">Ausgestorbene Wortbildungsmorpheme werden nach Möglichkeit durch entsprechende nhd. Bildungen ersetzt. (siehe Beispieltabelle am Schluss)</td></tr>
<tr><td>dipl</td><td>clean</td><td>norm</td></tr>
<tr><td>halben (Postposition)

stachelecht</td><td>halben

stachelecht</td><td>halber (allenthalben ist ein Adverb und bleibt allenthalben)
stachelig</td></tr>
<tr><td rowspan="3">3.7</td><td colspan="3">Flexionslose Adjektive, die im Neuhochdeutschen eine overte Flexion enthalten, werden in der norm flektiert, ohne dass der Kasus geändert wird.</td></tr>
<tr><td>das kal haupt</td><td>das kal haupt</td><td>das kahle Haupt</td></tr>
<tr><td>bei ſchweinem fleiſch</td><td>bei schweinem fleisch</td><td>bei schweinenem Fleisch</td></tr>
<tr><td rowspan="5">3.8</td><td colspan="3">Intervenieren innerhalb eines Wortes Sonderzeichen, dann werden die Sonderzeichen in der norm-Ebene ignoriert und das unterbrochene Wort zusammengeschrieben.</td></tr>
<tr><td>dipl</td><td>clean</td><td>norm</td></tr>
<tr><td> 
<lb></td><td>Glo ✠
⊙
ver
Grass</td><td>Glover
Grass</td></tr>
<tr><td colspan="3">Text hinter (geschweiften) Klammern, wird nach dem umklammerten Textabschnitt eingefügt.</td></tr>
<tr><td colspan="3">dipl</td></tr>
<tr><td colspan="3"> 
Wermuth Wermuth , gemeiner Wermuth , groſzer ſ. Wer= muthbeifuſz .</td></tr>
</table>

	Graphematisch wird grundsätzlich nach dem Duden normalisiert.

Fußnoten und Marginalien

Der Text der **Marginalie** wird immer am Anfang des jeweiligen Absatzes, in dem/neben dem sie realisiert ist, unabhängig davon, ob so mehrere Marginalien hintereinander oder/und auf der nachfolgenden Seite realisiert werden müssen, in die dipl-Ebene/Transkription aufgenommen.
Der Text der **Fußnote** wird immer am Ende des Absatzes, in dem sie eingepflegt ist, unabhängig davon, ob sich so die Fußnoten sammeln oder erst auf der nachfolgenden Seite realisiert werden müssten, in die dipl-Ebene/Transkription aufgenommen.

Zeichentabelle

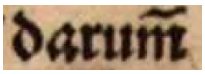
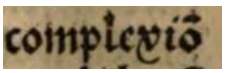
Beschreibung	Zeichen	Eingabe über die Tastatur
Schräger Dopppelbindestrich	⸗	2E17
Langes kleines „s“	ſ	017F Alt s
Tilde	̃	0303
Ligatur ae	æ	00E6
Ligatur AE	Æ	00C6
Ligatur oe	œ	0153
Ligatur OE	Œ	0152
Akut	´	dead key + <Vokal>
Gravis	`	shift+dead key+ <Vokal>
Cedille klein	ç	00E7
Cedille groß	Ç	00C7
Superskribiertes „e“	ᵉ	0364
Kreis	å	030A
y mit Punkten	ÿ	00FF
Absatzzeichen	¶	00B6
Abkürzung „der“	ð	AltGr+d
Häkchen über Zeichen	v̊	0309
Zeichen für lat. „recipe“	℞	211E
Zeichen für lat. „libra“	℔	2114
Zeichen für lat. „uncia“	℥	2125
Zeichen für lat. „scrupel“	℈	2108
3 Kreise	∴	2234
Kreis mit Punkt “einjährig”	⊙	2299
Latin Small Letter Rum Rotunda	ꝵ	A75D
Tironisches “et”	⁊	204A

Tabelle SEQ Tabelle * ARABIC 2: Zeichentabelle

Tabelle: Normalisierungsbeispiele

	clean	**norm**
1	**-icht/-echt/-acht/-lote**	**-(l)ich/-(l)ig**
	ästicht	ästig
	knöpfflicht	knöpfflich
	haarecht	haarig
	holzecht	holzig
	schattecht	schattig
	langlecht	länglich
	laulecht	laulig
	weißlecht	weißlich
	rohtlecht	rötlich
	stachelecht	stachelig
	gelblote	gelblich
2	**deutlich erkennbare Ortsnamen werden als nhd. Form normalisiert**	
	Franckfort	Frankfurt
	Muttetz	Muttenz
3	**-lich-Adverbien werden durch nhd. lich-lose Formen ersetzt**	
	kräftiglich	kräftig
	wunderbarlich	wunderbar
4	**für-**	**vor-**
	fürnehmlich	*vornehmlich*
5	**-für**	**-vor**
	herfür	hervor
6	**on- / ohne entspricht dem Präfix un-**	
	on- / ohne	un-
7	**fnhd. *etwan***	
	etwan	nhd. etwa (Adverb) nhd. etwas (Pronomen)
8	**-erin**	**flektierte nhd. Form**
	in ein erhabnes Steinerin Grab gelegt	steinernes
9	***-en***	***-Ø***
	endlichen	*endlich*
	sonsten	*sonst*

	selbsten	*selbst*
10	***-Ø***	***-en***
	ob	*oben*
	beizeit	*beizeiten*
11	**(Nicht)Normaliserung von Kasus**	
	mit Normalisierung: Man trinke des wassers	mit Normalisierung von Kasus: *Man trinke das Wasser*
	Man trinke des wassers	ohne Normalisierung von Kasus: *Man trinke des Wassers*
12	**Vom Nhd. abweichende starke oder schwache Verbalflexion bzw. Rückumlaut wird dem nhd. Stand angepasst**	
	gennent *gebauen*	*genannt* *gebaut*
13	**Auseinanderschreibung**	
	obgenannt *hiebevor*	*oben genannt* *hier bevor*
14	**ggf. ge-Präfix in Partizipien ergänzen**	
	funden	gefunden
	kommen	gekommen
	worden	geworden
15	**Flexionsformen der Pronomina *der, die, das* an nhd. Formen anpassen**	
	der	derer
	dero	derer
	des	dessen
16	**fnhd. *wann/wenn***	nhd. wann (Interrogativpron) nhd. wenn (Subjunktion)
	wann wenn es einen Nebensatz einleitet → keine V2 Stellung, sondern VLetzt (auch VLetzt mit Nachfeldbesetzung: *wenn man Köl- und Haselbaum pflanzt zu Weinreben*)	wenn
	wann wenn es einen Hauptsatz einleitet → V2 und übersetzbar it nhd. ‚denn‘	wann
	wenn	wann

	wenn es einen Hauptsatz einleitet → V2 und übersetzbar it nhd. ‚denn‘	
	wenn wenn es einen Nebensatz einleitet → keine V2 Stellung, sondern VLetzt (auch VLetzt mit Nachfeldbesetzung: *wenn man Köl- und Haselbaum pflanzt zu Weinreben*)	wenn
17	**fnhd. söllen/wöllen sind dialektale Formen**	
	söllen	sollen
	wöllen	wollen
18	**Eigennamen/Fremdwörter in Komposita werden getrennt geschrieben**	
	St. Anthonius Feuer	Sankt Anthonius Feuer
	Cassia fistel	Cassia Fistel
19	**Abkürzung frequenter Wörter und Endungen**	
	unn (sic!)	und
	umm (sic!)	um
	 darumm (sic!)	darum
	 complexiom (sic!)	Komplexion
20	**Einzelentscheidungen**	
	benommen	genommen
	Beweisung	Beweisung
	brauch (= Gebrauch)	Brauch
	dannenher	daher
	dennocht	dennoch
	der selbe	der selbe
	derselbe	derselbe
	derowegen	deswegen
	entbeut	entbietet
	Epff	Epff
	fahen	fangen
	gel	gel
	gepreist	gepriesen

gleich(e)sfalls	gleichfalls
guldin	golden
halb	halber
harm	Harn
harmen	harnen
hauffecht	häufig
leichtlich	leicht
liebnusz	Liebnis
Mannen	Männern
mehrer teils	mehrenteils
morgen (= morgens)	morgens
namlichen	nämlich
Nutz	Nutzen
pestnachen	Pastinaken
samlen	sammeln
schlahen	schlagen
sonders	sonders
(wegen der) Viele	Viele
viereckt	viereckig
vor	vorher
wehtum/wehtumb	Wehtun
wofer	wofern
wurz	wurzel
yedickest	gedihest (Lexer: ‚oft')
yetliches	jegliches
zeuchen	ziehen
wa (als dialektale Form)	wo
Vättern	Väter
nießte	neuste

In Verbindung mit Excel benutzte Tools

Excel-Makro
SearchAndMerge.bas: https://hu.berlin/search-and-merge
Das Makro SearchAndMerge.bas sucht in einem markierten Bereich nach leeren Zellen (kleiner gleich 300 leere Zellen nacheinander) und vereinigt diese mit der letzten Vorgängerzelle, die einen Wert enthält.

Python-Skript
clean-skript_V3.py: https://hu.berlin/clean-script-v3
Die neuen Texte ab Version 5 stellten neue Anforderungen an die <clean>-Ebene, neue Zeichenersetzungen wurden nötig sowie ein komplett neuer Umgang mit Nasalstrichen. Die vorigen Texte ließen Zeichenersetzungen der Nasalstriche zu den jeweiligen Nasalen durch Kontextbetrachtung zu, während die neuen Texte viele von diesen alten Regularitäten nicht mehr erfüllen. Deshalb wird ab der zweiten Version des clean-Skriptes anders mit solchen Nasalstrichen verfahren. Statt den Kontext zu betrachten und eine eindeutige Entscheidung zugunsten eines Nasals zu treffen, werden nun alle in Frage kommenden Zeichenersetzungen berücksichtigt und die möglichen Token werden durch „|“ getrennt dargestellt.

Bei der weiteren Bearbeitung in Excel kann es nach Anwendung des Clean-Skripts zu ungewollten automatischen Formatierungen kommen, ÜBERPRÜFEN!!!
1/2 wurde durch Excel zu *01.Februar.* Das normalisierte *falsch* wurde durch Excel zum logischen Operator *FALSCH.*