

Historische Korpuslinguistik (a.k.a. „Sprachgeschichte“)

Blockseminar, Universität Siegen, 18.–20. Februar 2020

Hinweis: Der Seminarplan dient der nur groben Orientierung – gerne passe ich die Inhalte spontan an Ihre Fragen, Wünsche, Anregungen etc. an. Wenn Sie nicht dafür sorgen, dass es zu groben Abweichungen vom Plan kommt, tue ich es ganz sicher...

(Auch sind im Seminarplan außer den Mittagspausen noch keine Pausen vermerkt. Kaffee-, Zigaretten- oder sonstige abhängigkeits- und bedürfnisstillende Pausen können wir spontan einschieben.)

Dienstag, 18.02.2020

09.00–09.30	Einführung und Organisatorisches
09.30–10.30	Korpuslinguistik und Sprachgeschichte Korpuslinguistik: Wieso, weshalb, warum? Exkurs: Prinzipien der Korpuserstellung Ein kurzer Rundgang durch die deutsche Korpuslandschaft Historische Korpora
10.30–12.00	Grundbegriffe am praktischen Beispiel Korpus und Konkordanz, Types und Tokens, ...
12.00–13.00	Mittagspause
13.00–14.30	Wie finde ich, was ich suche? (1) Literaturempfehlungen und Ressourcen; einschlägige Software Korpusabfrageprogramme und Abfragesyntax Reguläre Ausdrücke
14.30–16.00	Wie finde ich, was ich suche? (2) Übungen zu regulären Ausdrücken
16.00–17.30	Hinter den Kulissen eines Korpus (1) Encoding Hell Hinter den Kulissen – GerManC-Korpus GerManC: Praktische Übung
17.30–18.00	Puffer

Mittwoch, 19.02.2020

09.00–11.00	Deutsches Textarchiv und DWDS Welche Korpora bietet DWDS? Abfragesyntax Praktische Übung mit DWDS
11.00–12.00	Praktische Übungen in Excel Von der Konkordanz zur Analyse
12.00–13.00	Mittagspause
13.00–14.30	Deutsch Diachron Digital und ANNIS Welche Korpora sind über ANNIS verfügbar? Einstieg in ANNIS ANNIS-Suchabfragesyntax
14.30–16.00	ANNIS exportieren Exportmöglichkeiten von ANNIS

	Wie gewinne ich eine „echte“ Konkordanz aus ANNIS-Daten? Praktische Übungen mit DDD, Bonner Frühneuhochdeutschkorpus und ANNIS
16.00–17.30	Hinter den Kulissen eines Korpus (2) Eigene Skripts als Alternative zu Korpusabfrageprogrammen am Beispiel von REM
17.30–18.00	Puffer

Donnerstag, 20.02.2020

09.00–10.00	Die Mutter aller (deutschsprachigen) Korpora: DeReKo Einführung ins COSMAS II-Interface Aufbau des Deutschen Referenzkorpus COSMAS II-Abfragesyntax Praktische Übungen mit COSMAS II
10.00 – 12.30	Webkorpora und Zweifelsfälle Das Web als Korpus Wacky und COW – Möglichkeiten und Grenzen SketchEngine und NoSketchEngine Google ngram Viewer
12.30–13.30	Mittagspause
13.30–15.00	Statistische Auswertung von Korpusdaten Grundwissen Statistik und quantitative Methoden
15.00–16.00	Letzte Worte (Historische) Korpuslinguistik im 21. Jahrhundert: „Best Practice“ und Perspektiven