

EMLex-Aufbaumodul A6

1	Modulbezeichnung	Aufbaumodul A6 – Hauptseminar: Computerlexikographie Advanced module A6 – Advanced seminar: Computational Lexicography	Gesamtzahl ECTS- Punkte
2	Lehrveranstaltungen	2 SWS, Blockseminar: 5-8 Tage, 24-30 Unterrichtsstunden	5
3	Lehrende	Ulrich Heid / Stefan Evert / Besim Kabashi	

4	Modulverantwortliche/-r	Stefan Evert	
5	Inhalt	<p>Grundlagen der Korpuslinguistik</p> <ul style="list-style-type: none"> • Ansätze und Methoden zur Auswertung von Korpora • Lexikographische Anwendungen von Korpusdaten • Korpus-Typologie, Übersicht wichtiger verfügbarer Korpora • Korpusdesign, Repräsentativität, Textquellen, Metadaten <p>Erstellung eigener Korpora</p> <ul style="list-style-type: none"> • Korpuserstellung aus Online-Quellen per Web-Scraping etc. • Bereinigung und Normalisierung, Gewinnung von Metadaten • Repräsentations- und Austauschformate (Standards) • Online- und Offline-Werkzeuge für Web-Scraping etc. • Automatische linguistische Annotation (Wortarten = POS, Lemmatisierung, Eigennamen, syntaktische Analyse, ...) • Online- und Offline-Werkzeuge für die automatische linguistische Annotation <p>Suche in Korpora</p> <ul style="list-style-type: none"> • Reguläre Ausdrücke • Zeichensätze, Unicode-Standard • CQP-Anfragesprache zur Suche nach lexikogramatischen Mustern • Übungen mit SketchEngine und CQPweb <p>Quantitative Auswertungen</p> <ul style="list-style-type: none"> • Häufigkeitslisten und Metadaten-Verteilung • Kollokationen und Word Sketches • Keyword-Analyse • Lexikographische Interpretation der Ergebnisse • Grundprinzipien der statistischen Inferenz <p>Reproduzierbarkeit</p> <ul style="list-style-type: none"> • Forschungsmethodik und Dokumentation <p>Datenmanagement, Nachhaltigkeit von Korpora und Ergebnissen</p>	
6	Lernziele und Kompetenzen	Die Studierenden sollen in die Lage versetzt werden,	

		<ul style="list-style-type: none"> • ihren Bedarf an Korpusdaten für ein lexikographisches Projekt zu formulieren und ein geeignetes Korpus zu konzipieren; • ein solches Korpus aus Webseiten oder anderen Quellen zu erstellen; • das Korpus mit Hilfe automatischer Werkzeuge linguistisch zu annotieren; • im Korpus mit regulären Ausdrücken und komplexeren Suchanfragen nach lexikogramatischen Mustern zu suchen; • quantitative Methoden wie Kollokations- und Keyword-Analysen anzuwenden und die Ergebnisse in angemessener Weise zu interpretieren; • ihre Ergebnisse in Kurzvorträgen zu vermitteln; • wissenschaftliche Diskussionen über technische und methodologische Aspekte der Korpusarbeit zu führen; und ihre Korpora und Auswertungsergebnisse zu dokumentieren und zu archivieren.
7	Voraussetzungen für die Teilnahme	a) 25 ECTS-Punkte aus dem ersten Semester
8	Einpassung in Studienverlaufsplan	Wahlmodul im zweiten Semester (6 bzw. 7 von 10)
9	Verwendbarkeit des Moduls	für den Masterstudiengang EMLex
10	Studien- und Prüfungsleistungen	nach Wahl der Studierenden (b wird empfohlen): a) 90-minütige Abschlussklausur zu den Seminarinhalten <i>oder</i> b) Präsentation (Kursprojekt) plus kurze Hausarbeit (ca. 10 Seiten) <i>oder</i> längere Hausarbeit (15-20 Seiten)
11	Berechnung der Modulnote	100% der Modulabschlussprüfung
12	Turnus des Angebots	jährlich, nur im Sommersemester
13	Wiederholung der Prüfungen	1
14	Arbeitsaufwand in Zeitstunden	Präsenzzeit: 5-8 Tage, 24-30 Unterrichtsstunden Eigenleistung: 120 Stunden
15	Dauer des Moduls	Blockveranstaltung (Termin und Dauer nach besonderer Bekanntgabe)
16	Unterrichts- und Prüfungssprache	Englisch oder Deutsch
17	Literaturhinweise	HSK 5.4, Kap. XVIII + XIX
18	letzte Änderung	März 2022