

## EMLex-Aufbaumodul A6

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

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

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