

Master of Science Digital Humanities

Akademischer Grad	Modulnummer	Modulform
Master of Science	10-DIH-1001	Wahlpflicht

Modultitel Einführung in linguistische Annotation und XML Technologien

Modultitel (englisch) Introduction to Linguistic Annotation and XML Technologies

Empfohlen für: 2. Semester

Verantwortlich Lehrstuhl für Digital Humanities

Dauer 1 Semester

Modulturnus jedes Sommersemester

Lehrformen

- Vorlesung "Einführung in linguistische Annotation und XML Technologien" (2 SWS) = 30 h Präsenzzeit und 120 h Selbststudium = 150 h
- Praktikum "Praktikum" (0 SWS) = 0 h Präsenzzeit und 150 h Selbststudium = 150 h

Arbeitsaufwand 10 LP = 300 Arbeitsstunden (Workload)

Verwendbarkeit • M.Sc. Digital Humanities

Ziele

Nach der aktiven Teilnahme am Modul "Einführung in linguistische Annotation und XML Technologien" sind die Studierende in der Lage, die Hauptkomponenten der linguistischen Annotation zu erkennen und zu definieren. Dazu gehören die Kenntnis der Konzepte Notationsschema, Annotation, Training- und Testphasen und Auswertung, sowie die Fähigkeit, den Unterschieden zwischen die wichtigsten XML Technologien (XML Schema, XQuery, XSLT) zu erklären.

Die Studierenden können einfache Texte gemäß den TEI-Richtlinien interpretieren und kodieren und Unicode-bezogene Probleme lösen. Sie können einfache XQuery/XSLT Skripte schreiben, um Texte abzufragen und zu transformieren.

Inhalt

Vorlesung:

- 1 Prinzipien der manuellen Annotation und Maschinenlernen
- 2 Text encoding: Dateien-Formate (TXT, XML, JSON), UNICODE, TEI (= Text Encoding Initiative)
- 3 XML und XML Schema foundations
- 4 XQuery: Extraktion und Manipulation von Informationen
- 5 XSLT: Transformation von XML Dokumenten

Praktikum:

- 1 Übungen zum Kodieren von Text, Annotation, POS Training, Taggers/Syntaktische Parsers sowie Query/Transformation von Dokumenten

Die Lehrveranstaltungen werden in englischer Sprache gehalten. Studien- und Prüfungsleistungen sind in englischer Sprache zu erbringen.

Teilnahmevoraussetzungen keine

Literaturangabe Hinweise zu Literaturangaben erfolgen in den Lehrveranstaltungen.

Vergabe von Leistungspunkten Leistungspunkte werden mit erfolgreichem Abschluss des Moduls vergeben. Näheres regelt die Prüfungsordnung.

Prüfungsleistungen und -vorleistungen

Modulprüfung: Elektronische Prüfung 90 Min., mit Wichtung: 1	
	Vorlesung "Einführung in linguistische Annotation und XML Technologien" (2SWS)
	Praktikum "Praktikum" (0SWS)