Modulbezeichnung	**Datenbank- und Suchtechnologien**
Modulverantwortliche(r)	Prof. Dr. S. Maneth
Modulart	Pflicht/Wahl ☒ Wahlpflicht ☐
Spezialisierungsbereich	
Dauer des Moduls	1 Semester
Kreditpunkte	6 CP
Arbeitsaufwand	Berechnung des Workloads Präsenz 56 h Übungsbetrieb/Prüfungsvorbereitung 124 h Summe 180 h
Turnus des Moduls	unregelmäßig
Voraussetzung für die Teilnahme	Keine ☐ Folgende Formale Voraussetzungen: Keine
Lehr- und Lernformen	Seminar ☐ Vorlesung ☒ Tutorium ☒ Praktikum ☐ Projekt ☐
Lernziele	• In der Terminologie der Fachgebiete Datenbanken und Information Retrieval kommunizieren koennen, Teilprobleme und Methodiken des Gebietes unterscheiden und klassifizieren können. • Exakte Syntax von XML und von Document Type Definitions (DTDs) verstehen. • Dokumente mittels XML Parsers (SAX und DOM) einlesen und bearbeiten können. • Verständnis von Datenbanknormformen und deren Zusammenhänge. • Grundlegende SQL queries formulieren können • Grundzüge der Volltextsuche verstehen, insbesondere TF-IDF ranking und invertierte files. • Genaues Verständnis von online-Textsuche Algorithmen (KMP, Boyer-Moore, Horspool). • Genaues Verständnis von offline (indizierter) Textsuche mittels Suffix Trees und Arrays und Burrows Wheeler Transform. • Grundzüge der XML Abfragesprache XPath verstehen und anwenden können.

Lerninhalte	• Introduction, Basics of XML • Document Type Definitions (DTDs) • DTDs (regular expressions) and DOM • SAX Parsing, Entity Relationship Model • ER Model, Normal Forms • Normal Forms • Simple SQL queries • SQL and beyond • Spatial queries and indexes in MySQL • Text indexing (inverted files), scoring • TFIDF Scoring, Lucene • Online Text Matching • KMP, Boyer-Moore, Horspool • Indexed String Search • Suffix Trees and Suffix Arrays • Suffix Array, Burrows-Wheeler Transform • XPath • XSLT
Prüfungsformen	i.d.R. Übungsaufgaben und Programmieraufgaben
Literatur	