

Modulverantwortliche/r Prof. Dr. Matthias Hemmje

Dauer des Moduls  
ein Semester

ECTS  
10

Workload  
300 Stunden

Häufigkeit  
in jedem Wintersemester

Lehrveranstaltung(en) Information Retrieval

Detaillierter Zeitaufwand Bearbeitung der Lektionen (7 mal 20 Stunden): 140 Stunden  
Einüben der Inhalte (7 mal 15 Stunden): 105 Stunden  
Wiederholung und Prüfungsvorbereitung: 55 Stunden

Qualifikationsziele Die Studierenden beherrschen die wichtigsten allgemeinen Begriffe, Modelle und Methoden des Information Retrieval, können diese innerhalb exemplarischer Technologien und Anwendungen aus dem Bereich Multimedia (Text, Bild, Video, Ton) zuordnen sowie innerhalb eigener Entwürfe und Modellierungen von Information Retrieval-Anwendungen in Benutzungsschnittstellen zu Informationssystemen anwenden.

Inhalte Beständig, und in zunehmendem Maße, werden Informationen direkt in digitaler Form erzeugt oder nachträglich in ein digitales Format überführt. Ein Grund dafür ist die schnelle und einfache Verarbeitung und eine damit einhergehende bessere Wiederverwendbarkeit. Einen umfangreichen digitalen Datenbestand jedoch manuell und gezielt nach einer bestimmten Information zu durchsuchen ist ab einer bestimmten Menge an Daten nicht mehr effektiv möglich und der tatsächliche Nutzen des Bestands damit zumindest fraglich. Ein plakatives Beispiel für einen multimedialen Datenbestand ist das Internet, welches massive Mengen an digitalen Daten vorhält. Wohlbekannte Suchmaschinen helfen hier dem Suchenden, um sich in diesem Bestand zurechtzufinden. Große Datenbestände entstehen jedoch auch in spezielleren Bereichen, wie z. B. in Behörden, Krankenhäusern oder Verlagen. Auch hier muss ein effektives Auffinden gesuchter Informationen gewährleistet werden. Die Forschung im Umfeld des Information Retrieval (IR) und des Multimedia Information Retrieval (MMIR) befasst sich daher mit der Modellierung und Umsetzung von Anwendungen, die automatisiert digitale Datenbestände für den einfachen Zugriff und die Nachnutzung aufbereiten. Die Forschung an effektiven IR- und MMIR-Verfahren ist hinreichend komplex, und obwohl das IR auf eine lange Historie zurückblickt, sind insbesondere mit Hinblick auf anwachsende Datenmengen mit zunehmend heterogener und verteilter Natur Fragestellungen offengeblieben und neue Anforderungen hinzugekommen. Die Verarbeitung von multimedialen Inhalten jeglicher Art stellt weitere Herausforderungen für MMIR dar. In diesem Kontext spielen Multimedia-Merkmale (sog. Features), deren Extraktion, Repräsentation und Nutzung für IR-Prozesse eine zentrale Rolle. Diese Lehrveranstaltung befasst sich zunächst mit klassischen Themen des IR in Dokumentdatenbeständen, um in die grundlegenden Eigenschaften einzuführen. Darunter fallen Themen wie die Indexierung von Text und Verfahren zu Gewichtungen von Indexeinheiten, die Einführung etablierter IR-Klassen und Modelle sowie Verfahren zur Evaluation von IR-Verfahren. Über die klassischen IR-Verfahren hinaus widmet sich diese Lehrveranstaltung dem Transfer dieser Verfahren auf MMIR und erweiterten Konzepten aus diesem Kontext, wie der semantischen Suche und verteilten MMIR-Verfahren.

Inhaltliche Voraussetzung Kenntnisse aus dem Modul 63413 "Dokumenten- und Wissensmanagement im Internet".  
Empfohlen: Modul 6414 "Multimedialinformationssysteme"

Lehr- und Betreuungsformen Lehrveranstaltungsmaterial  
internetgestütztes Diskussionsforum  
Einsendeaufgaben mit Korrektur und/oder Musterlösung

Anmerkung

Formale Voraussetzung keine

Verwendung des Moduls M.Sc. Data Science  
M.Sc. Informatik  
M.Sc. Praktische Informatik  
M.Sc. Wirtschaftsinformatik

| <b>Prüfungsformen</b>   |      | Art der Prüfungsleistung | Voraussetzung |
|-------------------------|------|--------------------------|---------------|
| Prüfung                 |      | benotete zweistündige    |               |
| Stellenwert<br>der Note | 1/12 | Prüfungsklausur          |               |