63413 Dokumenten- und Wissensmanagement im Internet

Modulverantwortliche/r: Prof. Dr. Matthias Hemmje

Dauer des Moduls	ECTS	Workload	Häufigkeit
ein Semester	10	300 Stunden	in jedem Semester

Lehrveranstaltung(en): Dokumenten- und Wissensmanagement im Internet

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:
Das hauptsächliche Ziel dieser Lehrveranstaltung ist der Erwerb von praktischer Kompetenz und Handlungsfähigkeit.
Die Studierenden sind nach Abschluss der Lehrveranstaltung in der Lage, das Modell der strukturierten Dokumente zu beschreiben und zu beurteilen. Sie können die wichtigsten Begriffe, Modelle und Methoden der Anwendung im Kontext von wissensbasierten Informationssystemanwendungen für das Internet formulieren. Teilnehmende der Lehrveranstaltung erhalten Kenntnisse über HTML, CSS und JavaScript und verstehen es, entsprechenden Code zu analysieren und eigenständig zu erstellen.
Weiterhin erschließen sich die Studierenden die grundlegenden Strukturen von XML-Dokumenten, sowie den zugehörigen Schemavarianten DTD und XSD, und können auf dieser Basis eigene Codebeispiele generieren.
Die Studierenden erarbeiten die grundlegende Funktionsweise und den Aufbau von XSLT Templates, sowie Suchanfragen an XML-Informationsquellen mittels XQuery. Sie sind in der Lage, Anfragen an XML-Dokumente mit XQuery selbständig zu formulieren.
Als Basis des Themengebietes Semantisches Web beschäftigen sich die Studierenden im Folgenden mit dem Begriff der Semantik und dem sogenannten Ressource Description Framework (RDF).
Abschließend wird den Studierenden ein Einblick in Aufbau und Anwendungsgebiete von Ontologiesprachen für das Internet vermittelt. Sie entwickeln ein Verständnis für die Rolle von Ontologien bei der Informationsintegration.

Inhalte:
Lektion 1 – Das Modell der strukturierten Dokumente
Lektion 2 – Einführung in HTML, CSS und JavaScript
Lektion 3 – Einführung in XML, DTD und XSD
Lektion 4 – Einführung in XSLT und XQuery
Lektion 5 – Semantik und Semantisches Web
Lektion 6 – Semantische Integration
Lektion 7 – Ontologiesprachen für das Internet

In dieser Lehrveranstaltung werden allgemeine Formate für Dokumente sowie für strukturierte Dokumente, die im Internet verfügbar sind, vorgestellt. Es folgt eine detaillierte Betrachtung des Modells der strukturierten Dokumente, in der die Studierenden verschiedene Arten von Textdokumenten sowie Umsetzungsmöglichkeiten des Modells analysieren.

Weiter untersuchen sie verschiedene Modelle von Zeichenkodierungen, die den Inhalt von Dokumenten in maschinenlesbarer Form repräsentieren, sowie deren einzelnen Bestandteile.

Die Teilnehmenden der Lehrveranstaltung befassen sich mit der Historie der Markup-Sprachen und sind in der Lage, deren Entwicklung mit der Hypertext-Markup-Language (HTML) als Ausgangsbasis zu erläutern.
Die Anbindung von Stylesheets an HTML-Dokumente, die CSS-Syntax und die Verwendung von CSS-Selektoren werden ebenfalls dargestellt und können anhand von Praxisbeispielen übertragen und vertieft werden.

Im Anschluss wird im Ansatz auf JavaScript als Webprogrammiersprache eingegangen und die grundlegende Funktionsweise beleuchtet.

Die Studierenden setzen sich des Weiteren detailliert mit XML als einer Meta-Auszeichnungssprache, mit der andere Auszeichnungssprachen definiert werden können, sowie als Datenformat und insbesondere auch als Datenaustauschformat zwischen Anwendungen auseinander.
Die Studierenden lernen Dokumenttyp-Definitionen (kurz DTDs) und XML-Schema-Definitionen (kurz XSDs) kennen. Die Teilnehmenden der Lehrveranstaltung sind in der Lage, eigene Schemabeschreibungen für XML-Dokumente zu erstellen.

Anknüpfend an die Grundlagen von XML werden die Anwendungen XSLT und XQuery erarbeitet. Die Studierenden der Lehrveranstaltung können eigene XSLT-Templates, beispielsweise mit dem Ziel der selektiven Datenausgabe, erzeugen. Darauffolgend wird XQuery als Abfragesprache für XML-Dokumente eingeführt, mittels derer Suchanfragen auf XML-Datenquellen gestellt werden können.

Die Studierenden setzen sich weiterhin mit dem semantischen Web als Fortentwicklung des bestehenden WWWs auseinander. Betrachtet werden die Bereitstellung von Informationsquellen im Hinblick auf menschliche und auch maschinenlesbare Nutzung, sowie Probleme und Lösungsansätze bei der Etablierung eines Semantischen Netzes.

Die Studierenden analysieren das Ressource Description Framework (RDF) als ein Modell der Wissensrepräsentation und der Darstellung von Metadaten.

Als Ausgangspunkt des Ontologie-basierenden Informationsmanagements differenzieren die Studierenden zwischen drei Konzepten der semantischen Heterogenität. Der Begriff der Ontologie, sowie ihr Aufbau, auch im Vergleich zur Taxonomie wird näher betrachtet und ihre Rolle bei der Reduktion von semantischer Heterogenität durch Konzeptualisierungen und gemeinsam genutztes Vokabular beleuchtet.

Des Weiteren werden der Übersetzungsprozess und das Ontologie-Alignment betrachtet, wodurch die Interoperabilität zwischen semantischen (Web-)Applikationen gewährleistet werden soll. Die Begriffe Ontology Mapping und Ontology Merging sind nach der Bearbeitung ebenfalls bekannt und können kontrastiert werden.

Ergänzende Literatur:

P. Aiken, M. D. Allen, "XML in Data Management: Understanding and Applying Them Together", Morgan Kaufmann, June 7, 2004, ISBN-13: 978-0120455997
S. Abiteboul, P. Bunemann, D. Suciu: "Data on the Web: From Relations to Semistructured Data and XML" Morgan Kaufmann, October 12, 1999, ISBN-13 978-1558606227
G. Antoniou, F. van Harmelen: "A Semantic Web Primer", The MIT Press, April 1, 2004, ISBN-13: 978-0262012102
J. Davies, R. Studer, P. Warren (eds): "Semantic Web Technologies: Trends and Research in Ontology-based Systems", Wiley, July 11, 2006, ISBN-13: 978-0470025963
A. Sheth, M. Lytras: "Semantic Web-Based Information Systems: State-of-the-Art Applications" Cybertech Publishing, December 11, 2006, ISBN-13: 978-1599044279

Inhaltliche Voraussetzung

Programmierkenntnisse

Lehr- und Betreuungsformen

Lehrveranstaltungsmaterial

Einsendeaufgaben mit Korrektur und/oder Musterlösung

Internetgestütztes Diskussionsforum

Betreuung und Beratung durch Lehrende

Anmerkung

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

Prüfungsformen	Art der Prüfungsleistung	Voraussetzung
Prüfung	benotete zweistündige Prüfungsklausur	keine
Stellenwert der Note	s. PO	