

Datenintegration								Modulnummer:	
<i>Data Integration</i>									
Bachelor Pflicht/Wahl <input checked="" type="checkbox"/> Wahlpflicht <input type="checkbox"/> Wahl <input type="checkbox"/> Sonderfall <input type="checkbox"/>					Modulbereich: Pflicht				
Anzahl der SWS	V	UE	K	S	Prak.	Proj.	Σ	Kreditpunkte: 6	Turnus i.d.R. alle 2 Semester
	3	1	0	0	0	0	4		
Formale Voraussetzungen: Keine									
Inhaltliche Voraussetzungen: -									
Vorgesehenes Semester: ab 1. Semester									
Sprache: Deutsch									
Ziele: <ul style="list-style-type: none"> • In der Terminologie des Fachgebietes Datenbanken kommunizieren können, Teilprobleme und Methodiken des Gebietes unterscheiden und klassifizieren können. • Methoden der Schemaabbildungen zwischen relationalen Datenquellen verstehen und anwenden können. • Möglichkeiten zur Informationsextraktion (z.B. aus Web Seiten) verstehen und bewerten können. • Mechanismen zum Umgang mit fehlerhaften Daten verstehen und anwenden können. • Datenprovenienz innerhalb von relationalen Datenbanken modellieren können. 									
Inhalte: <ul style="list-style-type: none"> • Datenbankkonzepte: Integritätsbedingungen, konjunktive Anfragen (kA), Datalog Programme • Anfrageäquivalenz und Enthaltensein: kAs, Vereinigungen von kAs, kAs mit Negation, Multimengen und Aggregation • Anfrageauswertung auf Views: Bucket und MiniCon Algorithmus, Inversregel Algorithmus, Updates auf Views • Ontologie-Basierter Datenzugriff: Ontologiesprachen, DL-LITE, Rewriting versus Unfolding • Abbildungen von Schemata: Global-as-View (GAV), Local-as-View (LAV), Global-and-Local-as-View (GLAV), Tupelgenerierende Abhängigkeiten • Anfrageauswertung mit/ohne Bedingungen: GAV, LAV, und GLAV settings, Auswertung vs. Rewriting • Wrappers: manuelle und lernbasierte Konstruktion, Web-Wrappers, Systeme: Stalker, Poly und Lito • Umgang mit Unsicherheit und fehlerhaften Daten: zuverlässige Antworten, probabilistische Datenbanken, probabilistische Schemaabbildungen • Datenprovenienz: Annotationen und Datenabhaengigkeitsgraphen, der Provenienzhilbring 									
Unterlagen (Skripte, Literatur, Programme usw.): AnHai Doan, Alon Halevy, Zachary Ives: Principles of Data Integration, Morgan Kaufmann, 2012 Dan Such, Dan Olteanu, Christopher Re, Christoph Koch: Probabilistic Databases, Morgan & Claypool Publishers, 2011									
Form der Prüfung: i.d.R. Übungsaufgaben und Fachgespräch oder mündliche Prüfung									
Arbeitsaufwand		Präsenz		56 h		Übungsbetrieb/Prüfungsvorbereitung		124 h	
		Summe		180 h					
Lehrende: Prof. Dr. S. Maneth					Verantwortlich: Prof. Dr. S. Maneth				