

1.	Наслов на наставниот предмет	<b>Податочна фузија</b> Data fusion	
2.	Код	СБП-3-03	
3.	Студиска програма	магистерски студии по Информатички науки и компјутерско инженерство модул Содржински базирано пребарување	
4.	Организатор на студиската програма (единица, односно институт, катедра, оддел)	Факултет за информатички науки и компјутерско инженерство – ФИНКИ	
5.	Степен (прв, втор, трет циклус)	втор циклус	
6.	Академска година / семестар <b>2 / летен / задолжителен</b>	7. Број на ЕКТС кредити	<b>6</b>
8.	Наставник	Доц. д-р Игор Трајковски	
9.	Предуслови за запишување на предметот	Нема	
10.	Цели на предметната програма (компетенции): Во овој курс се изучуваат методи и стратегии за извлекување на информации од неструктурирани и полуструктурирани web-документи како и техниките за претставување на знаење и верување во компјутерските системи. Студентите ќе бидат запознаети со алгоритми за извлекување на имиња на ентитети и релации помеѓу овие ентитети од чист неструктуриран текст. Проблемот на извлекување на информации студентите ќе го разгледуваат од перспектива на машинско учење и имплементација на соодветни алгоритми. Главната цел за постоењето на знаењето и верувањето е расудувањето за и со нив. Во овој курс студентите ќе бидат запознаени и со начинот на работа на неколку системи кои користат формална логика, како во нив можат да се имплементираат процедури за расудување и како истите можат да се користат во проблемите на расудување.		
11.	Содржина на предметната програма: Вовед во проблематиката на извлекување на информации од текстуални податоци; Извлекување на имиња на ентитети (Name Entity Extraction, NER); Извлекување на релации; Извлекување на ентитети и релации од Wikipedia; Извлекување на онтологии од Wikipedia; Интегрирање на онтологии; Одговарање на прашања користејќи ја Wikipedia; KnowItAll и ReVerb; Лексикони и онтологии за претставување на здраворазумско знаење; SnePS; WordNet, Cys, FrameNet;		
12.	Методи на учење: Предавања поддржани со презентации преку слајдови, интерактивни предавања, вежби (користење на опрема и софтверски пакети), тимска работа, пример случаи, поканети гости предавачи, самостојна изработка и одбрана на проектна задача и семинарска работа, учење во електронско опкружување (форуми, консултации).		
13.	Вкупен расположив фонд на време	6 ECTS x 30 часа = 180 часа	
14.	Распределба на расположивото време	130 + 0 + 50 = 180 часа	
15.	Форми на наставните активности	15.1. Предавања- теоретска настава	130 часови
		15.2. Вежби (лабораториски, аудиториски), семинари, тимска работа	0 часови
16.	Други форми на активности	16.1. Проектни задачи	15 часови
		16.2. Самостојни задачи	15 часови
		16.3. Домашно учење	20 часови
17.	Начин на оценување		

17.1.	Тестови				65 бодови
17.2.	Семинарска работа/ проект ( презентација: писмена и усна)				25 бодови
17.3.	Активност и учество				10 бодови
18.	Критериуми за оценување (бодови/ оценка)	до 59 бода			5 (пет) (F)
		од 60 до 68 бода			6 (шест) (E)
		од 69 до 76 бода			7 (седум) (D)
		од 77 до 84 бода			8 (осум) (C)
		од 85 до 92 бода			9 (девет) (B)
		од 93 до 100 бода			10 (десет) (A)
19.	Услов за потпис и полагање на завршен испит	реализирани активности 15.1 и 15.2			
20.	Јазик на кој се изведува наставата	македонски и англиски			
21.	Метод на следење на квалитетот на наставата	механизам на интерна евалуација и анкети			
22.	Литература				
22.1.	Задолжителна литература				
	Ред. број	Автор	Наслов	Издавач	Година
	1.	Marie-Francine Moens	Information Extraction: Algorithms and Prospects in a Retrieval Context	Springer	2006
	2.	Sunita Sarawagi	Information Extraction	Now Publishers	2008
	3.	Stuart C. Shapiro	Knowledge Representation and Reasoning: Logics for Artificial Intelligence, Lecture Notes	University at Buffalo, The State University of New York Buffalo	2010
22.2.	Дополнителна литература				
	Ред. број	Автор	Наслов	Издавач	Година
	1.	Oren Etzioni & Co.	Unsupervised named-entity extraction from the web: an experimental study	Journal of Artificial Intelligence	2005
	2.	Cyc Corporation	Cyc 101 Tutorial	Cyc Corporation	2010
	3.		Publications Turing Centar, Investigating problems at the crossroads of natural language processing, data mining, Web search, and the Semantic Web. <a href="http://turing.cs.washington.edu/publications.htm">http://turing.cs.washington.edu/publications.htm</a>		