

1.	Наслов на наставниот предмет	Моделирање и управување на големи податоци Big Data Modeling and Management
2.	Код	КН-И-08
3.	Студиска програма	
4.	Организатор на студиската програма (единица, односно институт, катедра, оддел)	Факултет за информатички науки и компјутерско инженерство
5.	Степен (прв, втор, трет циклус)	втор циклус
6.	Академска година / семестар 5 / летен /	7. Број на ЕКТС кредити 6
8.	Наставник	вонр.проф. д-р. Горан Велинов, доц. д-р Ефтим Здравевски
9.	Предуслови за запишување на предметот	
10.	Цели на предметната програма (компетенции): Ќе бидат проучени трендовите на развој на традиционалните релациони (SQL) системи за управување со бази на податоци, складишта на податоци, како и концептите на NoSQL и NewSQL системи за управување со големи податоци. Ќе бидат разгледани концептите на складирање на податоците на различни мемориски медиуми за складирање. Ќе бидат проучени приодите за централно или дистрибуирано складирање, како и логичка организација по редици, колони, графови или документи. Ќе бидат проучени методите за партиционирање и индексирање на структурирани, (полу/не)структурирани и текстуални податоци. Ќе се обработат реални пристапи и решенија за надминување на предизвиците од аспект на моделирање, управување, имплементација и пуштање во продукција на системи со големи податоци. Студентите на крајот од предметот ќе знаат кои системи се најпогодни и кои чекори се потребни за воведување на системи за големи податоци во компаниите, како и кои се предизвиците со кои се соочуваат компаниите.	
11.	Содржина на предметната програма: Нов поглед на складиштата па податоци: концептуални, логички и физички модели; Концепти на податочни езера. Преглед на системите за управување со големи податоци. Моделирање на податоци кај системи со големи податоци: импликации на времето во моделирањето на податоците. Концепти бази на податоци организирани по колони (MonetDB, HBase, Cassandra, по клуч-вредност (DynamoDB, Riak), по документи (MongoDB, CouchDB), и во графови (Neo4j, OrientDB). Трансакциски и аналитички бази на податоци што работат во главна меморија. Алтернативни медијуми за складирање на податоци. Стратегии за индексирање и партиционирање и нивно влијание на скалабилноста и перформансите; Бази на податоци за индексирање на текст (Solr, Elasticsearch) Интегрирање на различни извори на податоци; Планирање на развој, капацитет и инфраструктура. Системи и алатки за анализа на статички големи податоци, како што се: Spark, Spark SQL, Hive, Pig, Tez и понови. Техники за справување со податочни потоци; Системи и алатки за анализа на динамички големи податоци како што се: Spark Streaming, Storm, Oozie, Sqoop, Flink и понови. Управување со пуштањето во производство на системите: очекувања, претпоставки, ризици, градење тимови	

	Стратегии и сценарија за миграции, безбедност и чување резервни копии на големи податоци.			
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	180		
14.	Распределба на расположивото време	30 + 30 + 45 + 30 + 45 = 180 часа		
15.	Форми на наставните активности	15.1.	Предавања- теоретска настава	30 часови
		15.2.	Вежби (лабораториски, аудиториски), семинари, тимска работа	30 часови
16.	Други форми на активности	16.1.	Проектни задачи	45 часови
		16.2.	Самостојни задачи	30 часови
		16.3.	Домашно учење	45 часови
17.	Начин на оценување			
	17.1.	Гестови	30 бодови	
	17.2.	Семинарска работа/ проект (презентација: писмена и усна)	50 бодови	
	17.3.	Активности и учење	20 бодови	
17.4.	Завршен испит	бодови		
18.	Критериуми за оценување (бодови/ оценка)	до 50 бода		5 (пет) (F)
		од 51 до 60 бода		6 (шест) (E)
		од 61 до 70 бода		7 (седум) (D)
		од 71 до 80 бода		8 (осум) (C)
		од 81 до 90 бода		9 (девет) (B)
		од 91 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		
21.	Метод на следење на квалитетот на наставата	механизам на интерна евалуација и анкети		
22.	Литература			
	22.1.	Задолжителна литература		

Ред.бр.	Автор	Наслов	Издавач	Година
1	Franz Faerber, Alfons Kemper, Per-Åke Larson, Justin Levandoski, Thomas Neumann and Andrew Pavlo	Main Memory Database Systems, Foundations and Trends in Databases	Now Publishers	2017
2	Daniel Abadi, Peter Boncz, Stavros Harizopoulos, Stratos Idreos and Samuel Madden	The Design and Implementation of Modern Column-Oriented Database Systems	Now Publishers	2013
3	García Márquez, Fausto Pedro, Lev, Benjamin	Big Data Management	Springer	2017
4	Corea, Francesco	Big Data Analytics: A Management Perspective	Springer	2016
5	Moshirpour, Mohammad, Far, Behrouz, Alhajj, Reda	Highlighting the Importance of Big Data Management and Analysis for Various Applications	Springer	2018
6	Sherif Sakr and Mohamed Gaber	Large Scale and Big Data: Processing and Management	CRC Press	2014
7	Shivnath Babu and Herodotos Herodotou	Massively Parallel Databases and MapReduce Systems	Now Publishers	2013

22.2. Дополнителна литература

Ред. број	Автор	Наслов	Издавач	Година